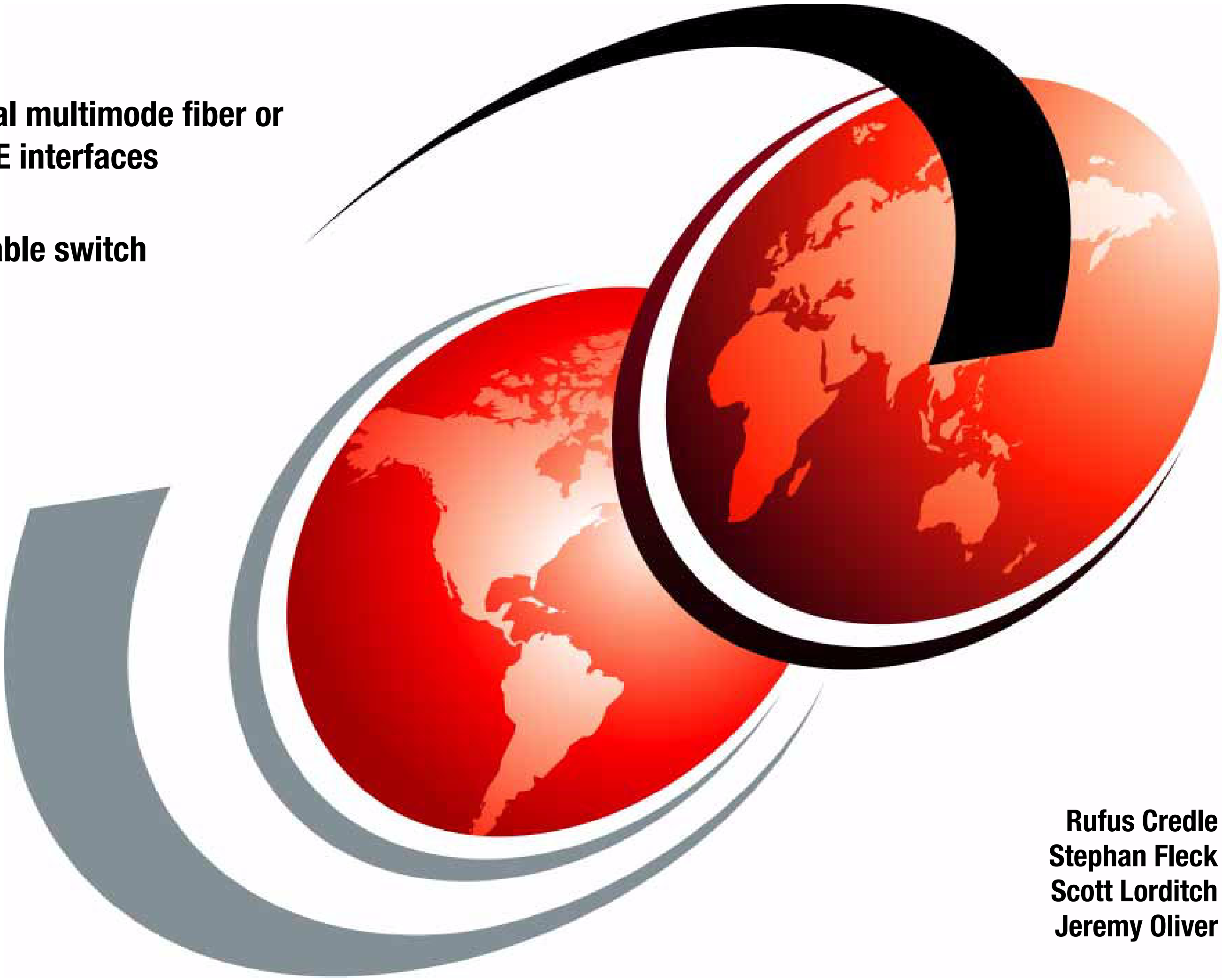
IBM eserver

Nortel Networks L2/3 Ethernet Switch Module for IBM eserver BladeCenter

- Full Layer 2 switching and Layer 3 routing
- Six external multimode fiber or copper GbE interfaces
- Hot pluggable switch modules



Rufus Credle
Stephan Fleck
Scott Lorditch
Jeremy Oliver

Redpaper

ibm.com/redbooks

International Technical Support Organization

Nortel Networks L2/3 Ethernet Switch Module for IBM @server BladeCenter

September 2005

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (September 2005)

This edition applies to Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM *@server* BladeCenter.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:
This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

@server®
@server®
Redbooks (logo) ™
Redbooks (logo)™
eServer™
xSeries®
AIX®
BladeCenter®
Domino®
Electronic Service Agent™
Enterprise Storage Server®
HelpCenter®
HelpWare®
IntelliStation®
IBM®
PartnerLink®
Redbooks™
ServerGuide™
Summit®
Tivoli®
TotalStorage®
WebSphere®

The following terms are trademarks of other companies:

Java, Sun, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Nortel Networks, the Nortel Networks logo, and the globemark design, and Alteon are trademarks of Nortel Networks.

The Extreme Networks logo, Alpine logo, BlackDiamond logo, Summit logos, and Extreme Turbodrive logo are trademarks of Extreme Networks.

Cisco, Cisco IOS, Cisco Systems, the Cisco Systems logo, EtherChannel are Registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

Other company, product, and service names may be trademarks or service marks of others.

Preface

This IBM® Redpaper positions the Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules for IBM @server BladeCenter and describes how its integrated switch options enable the consolidation of full Layer 2-3 LAN switching and routing capabilities. The Nortel Networks switch modules also provide an upgrade path to full Layer 4-7 services by including 4-7 switch intelligence.

This Redpaper serves as a Best Practices guide for implementing, configuring, and managing Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules for several network topologies. Our topology examples include Nortel Networks, Cisco Systems, and Extreme Networks network environments.

This Redpaper can help you to understand the Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules architecture. It demonstrates how to use specific tools to manage and administer switch module tasks. It also discusses the differences between Nortel Networks and Cisco Systems terminology.

The audience for this Redpaper is experienced systems and network administrators who want to integrate the Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules successfully into new and existing networks.

The team that wrote this Redpaper

This Redpaper was produced by a team of specialists from around the world working at the International Technical Support Organization (ITSO), Raleigh Center.

Rufus Credle is a Certified Consulting I/T Specialist and certified Professional Server Specialist at the ITSO, Raleigh Center. He conducts residencies and develops IBM Redbooks™ and Redpapers that discuss network operating systems, ERP solutions, voice technology, high availability and clustering solutions, Web application servers, pervasive computing, and IBM and OEM e-business applications, all running on IBM @server xSeries® and IBM @server BladeCenter® technology. Rufus's various positions during his IBM career have included assignments in administration and asset management, systems engineering, sales and marketing, and IT services. He holds a BS degree in business management from Saint Augustine's College. Rufus has been employed at IBM for 25 years.

Stephan Fleck is an IBM Accredited Senior IT Specialist at the EMEA ITS/TSS Networking Support Center. He has 12 years of experience in the networking area. Today, he provides EMEA-wide pre- and post-sales support. In addition to his technical skill, Stephan's expertise include project- and critsit-management. During his career, he has been active in product management, deploying new services for the field support group. Stephan is a Cisco Certified Internetwork Expert (CCIE #8301), and he holds a degree in Electrical Engineering from the Technical University Darmstadt, Germany. He has been employed at IBM for 11 years.

Scott Lorditch is a Sales Network Architect for the Blade Switching Server business unit of Nortel Networks. He develops designs and proposals for customers and potential customers of the Nortel Networks GbESM products for the IBM @server BladeCenter, including overall network architecture assessments. He also has developed several training and lab sessions for IBM technical and sales personnel and has provided field feedback to the product team. His background before working for Nortel includes almost 20 years working on networking, including electronic securities transfer projects for a major bank based in New York City, as

Senior Network Architect for a multi-national soft drink company, and as Product Manager for managed hosting services for a large telecommunications provider. He holds a BS in Operations Research with specialization in Computer Science from Cornell University.

Jeremy Oliver is a Staff Engineer with the System Validation and Storage group of xSeries Development. He has worked at IBM in Research Triangle Park, North Carolina for seven years. Jeremy's areas of expertise include developing experiments for testing new BladeCenter technologies as well as designing network and power infrastructures to handle test capacity, 10 Gb Ethernet, and operating systems. Jeremy holds a BS degree in Electrical Engineering from McNeese State University, Lake Charles, Louisiana. He also holds a MS degree in Electrical Engineering from North Carolina State University, Raleigh, with research in PHY technologies of computer networking.

Thanks to the following people for their contributions to this project:

Tamikia Barrows, Jeanne Tucker, Margaret Ticknor,
ITSO, Raleigh Center

Ishan Sehgal, BladeCenter Marketing Manager, Networking
IBM RTP

Paul Woodruff, General Manager of the Blade Server Switching business unit
Nortel Networks Santa Clara, CA

Shailesh Naik, Worldwide Director of the Sales Network Architect team
Nortel Networks Santa Clara, CA

Mark Davies, IBM Sales & Distribution xSeries FTSS
IBM Bermuda

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, Business Partners, or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this Redpaper or other Redbooks in one of the following ways:

- Use the online **Contact us** review redbook form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbook@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HQ7 Building 662
  P.O. Box 12195
  Research Triangle Park, NC 27709-2195

1

Executive summary

IBM and Nortel Networks are committed to collaborating on the design and development of server and networking technology to address customer requirements by establishing a joint development center. The Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter (Nortel Networks L2/3 GbESM and Nortel GbESM) represents a new height in this alliance.

The BladeCenter switch module offers BladeCenter customers Nortel's latest fiber and copper Gigabit Ethernet switching technology which is integrated into the BladeCenter chassis. It further enhances the BladeCenter value proposition by seamlessly interfacing to a customer's existing data network using six external multimode fiber or copper GbE interfaces.

When installed in the BladeCenter chassis, Nortel Networks L2/3 GbESM provides both full L2 switching and L3 routing capabilities and significant added value not found in commodity switching solutions. This value includes:

- VLAN tagging - 802.1Q
- Link Aggregation and LACP - 802.3ad and 802.3-2002
- Spanning Tree - 802.1D, 802.1w, 802.1s
- Routing Information Protocol - RFC1058 and RFC2453
- Open Shortest Path First (OSPF) - RFC1257, RFC2328, and others
- Virtual Router Redundancy Protocol (VRRP) - RFC 3768

Each Nortel Networks L2/3 GbESM provides one Gigabit per second Ethernet (GbE) connectivity to each of the 14 blade slots and six GbE uplink interfaces external to the IBM @server BladeCenter. The customer can install as few as one Nortel Networks L2/3 GbESM or as many as four Nortel Networks L2/3 GbESMs in one BladeCenter chassis. With four Nortel Networks L2/3 GbESMs installed, you can obtain 24 GbE uplink interfaces as well as 56 GbE internal switching capability. The flexibility of the Nortel Networks L2/3 GbESM allows you to address a variety of performance and redundancy needs.

The Nortel and IBM agreement to form a joint development center equips Nortel as it becomes an on demand company that can generate customized products for its network equipment marketplace. This ensures that your needs of high availability, scalability, security, and manageability are addressed. Combined with the integration of IBM Tivoli®, Nortel, and Cisco management products, these architectures bring higher value solutions with lower operational expense.

The Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter is an integral part of these solutions. With the Nortel Networks L2/3 GbESM, you have the investment protection and price performance of a solution behind which the world's leading server and networking companies stand.

2

IBM @server BladeCenter overview

IBM designed the IBM @server BladeCenter innovative modular technology, leadership density, and availability to help solve a multitude of real-world issues.

For organizations seeking server consolidation, the IBM @server BladeCenter centralizes servers for increased flexibility, ease of maintenance, reduced cost, and streamlined human resources. Companies that need to deploy new e-commerce and e-business applications can achieve speed while ensuring flexibility, scalability, and availability. For enterprise requirements such as file-and-print and collaboration, the IBM @server BladeCenter is designed to offer reliability, flexibility for growth, and cost effectiveness. In addition, clients with compute-intensive applications that need highly available clustering can use the IBM @server BladeCenter to help achieve high degrees of scalability and performance.

This chapter provides a high-level overview of the IBM @server BladeCenter product family.

2.1 The IBM @server BladeCenter product family

The IBM @server BladeCenter family of products features a modular design that integrates multiple computing resources into a cost-effective, high-density enclosure for a platform that:

- Reduces installation, deployment, and redeployment time
- Reduces administrative costs with our helpful management tools
- Achieves the highest levels of availability and reliability
- Provides XpandonDemand scale-out capability
- Reduces space and cooling requirements compared to 1U solutions

To understand more about how the Nortel Networks Layer 2/3 GbE Switch Module is designed to operate in the BladeCenter chassis, we suggest that you read the sections that follow which discuss the BladeCenter architecture. If you seek to know more about the IBM @server BladeCenter and its components, visit:

http://www.ibm.com/products/us/

Figure 2-1 on page 5 shows the IBM @server BladeCenter chassis, HS40, HS20, JS20, and LS20:

- IBM @server BladeCenter chassis

  The BladeCenter is a high-density blade solution that provides maximum performance, availability, and manageability for application serving, storage flexibility, and long-life investment protection.

- HS40

  HS40 is a 4-way blade server for high-performance enterprise applications requiring four-processor SMP capability. The BladeCenter chassis supports up to seven 4-way servers and is ideal for Enterprise Resource Planning (ERP) and database applications.

- HS20

  The IBM efficient 2-way blade server design offers high density without sacrificing server performance. Ideal for Domino®, Web server, Microsoft® Exchange, file and print, application server, and so on.

- JS20

  JS20 is a 2-way blade server for applications requiring 64-bit computing. Ideal for compute-intensive applications and transactional Web serving.

- LS20

  LS20 is a 2-way blade server running AMD Opteron processors. The LS20 delivers density without sacrificing processor performance or availability. For applications that are limited by memory performance, the LS20 might bring sizeable performance gains.

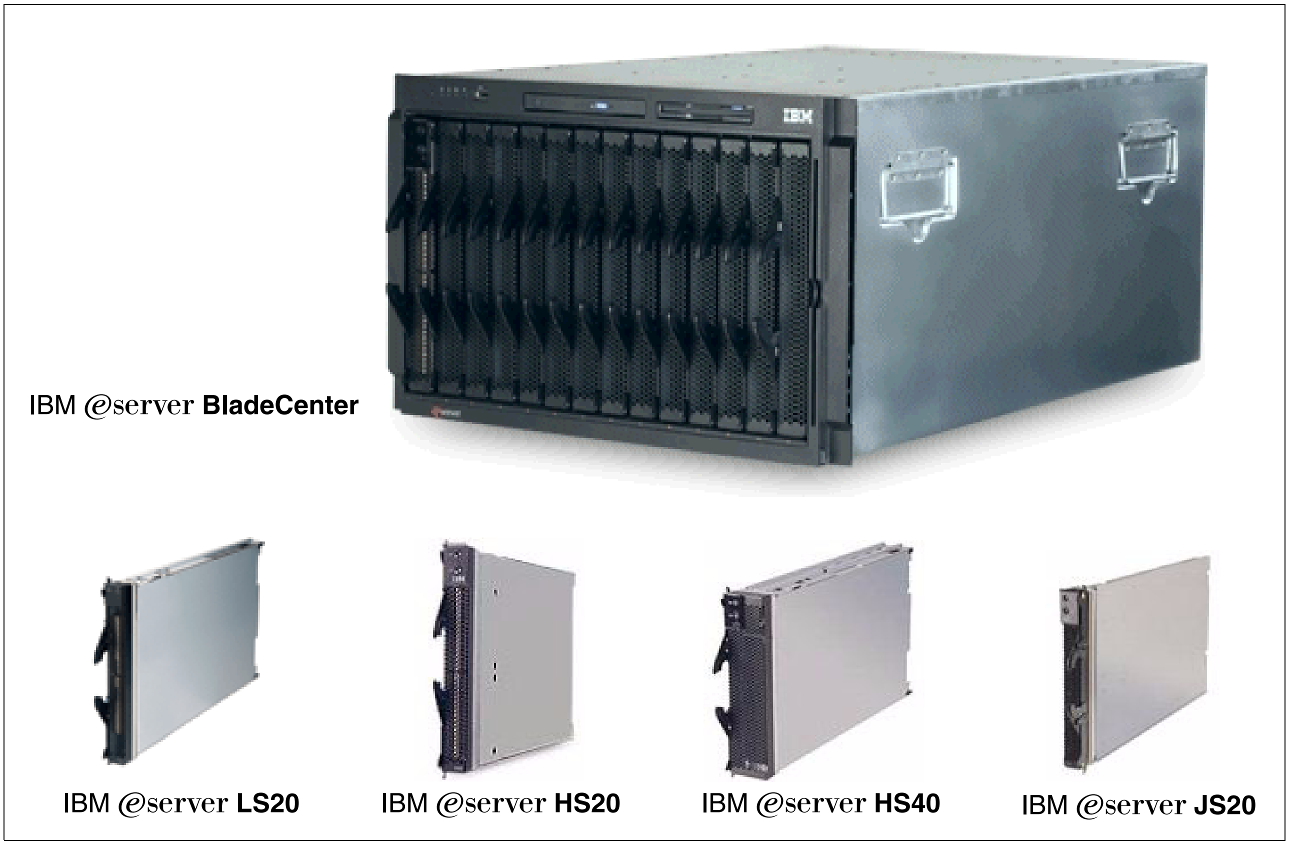


Figure 2-1 IBM @server BladeCenter and blade modules

Blade development is ongoing for the BladeCenter platform. Therefore, we suggest that you regularly visit the following Web site for the latest information about IBM @server BladeCenter:

http://www.ibm.com/servers/eserver/bladecenter/index.html

2.1.1 IBM @server BladeCenter storage solutions

IBM delivers a wide range of easy-to-install, high-capacity, tested storage products for the IBM @server BladeCenter to meet your demanding business needs. This enables you to choose from the array of IBM TotalStorage® storage solution products, which include:

- Fibre Channel products and Storage Area Networks
- Network Attached Storage
- Enterprise Storage Server®

IBM TotalStorage provides connected, protected, and complete storage solutions that are designed for your specific requirements, helping to make your storage environment easier to manage, helping to lower costs, and providing business efficiency and business continuity.

For more information about BladeCenter storage solutions, visit:

http://www.pc.ibm.com/us/eserver/xseries/storage.html

2.1.2 IBM @server BladeCenter system management

To get the most value from your IBM @server BladeCenter investment throughout its life cycle, you need smart, effective systems management which will keep your availability high and costs low.

Management foundation

IBM Director, our acclaimed industry standards-based workgroup software, delivers comprehensive management capability for IntelliStation®, ThinkCentre, ThinkPad, and IBM @server BladeCenter and xSeries hardware to help reduce costs and improve productivity. IBM Director is hardware that is designed for intelligent systems management. It offers the best tools in the industry and can save you time and money by increasing availability, tracking assets, optimizing performance, and enabling remote maintenance.

Advanced server management

This exclusive collection of software utilities provides advanced server management and maximum availability through the following components:

- Server Plus Pack
- Application Workload Manager
- Scalable Systems Manager
- Real-Time Diagnostics
- Electronic Service Agent™
- Tape Drive Management Assistant

For more information about advanced server management, see:

http://www-1.ibm.com/servers/eserver/xseries/systems_management/xseries_sm.html

Deployment and update management

IBM deployment tools help minimize the tedious work that can be involved in getting your servers and clients ready to run. These tools include:

- Remote Deployment Manager
- Software Distribution Premium Edition
- ServerGuide™
- ServerGuide Scripting Toolkit
- UpdateXpress

For more information about IBM @server BladeCenter deployment and update management, visit:

http://www.ibm.com/servers/eserver/xseries/systems_management/xseries_sm.html

2.2 IBM @server BladeCenter architecture

In this section, we look into the architectural design of the IBM @server BladeCenter chassis and its components.

2.2.1 The midplane

Figure 2-2 on page 7 illustrates the BladeCenter midplane. The midplane has two similar sections (upper and lower) that provide redundant functionality. The processor blades (blade servers) plug into the front of the midplane. All other major components plug into the rear of the midplane (for example, power modules, switch modules, and management modules). The processor blades have two connectors, one that is connected to the upper section and one that is connected to the lower section of the midplane. All other components plug into one section only (upper or lower). However, there is another matching component that can plug into the other midplane section for redundancy.

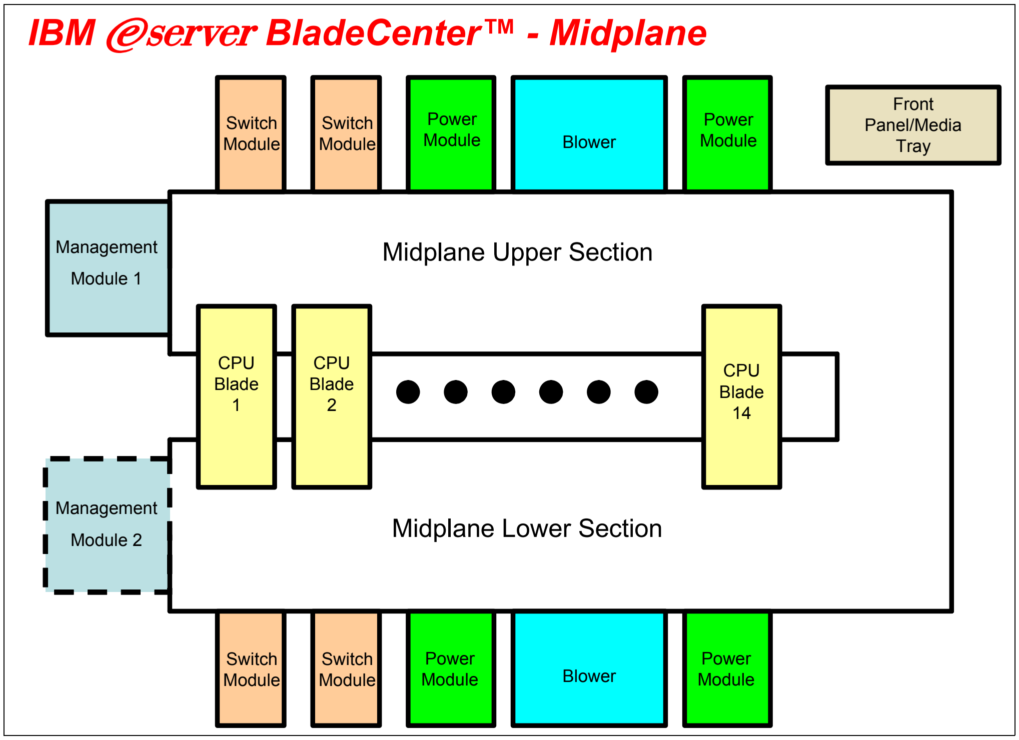


Figure 2-2 Midplane view

It should be noted that the upper and lower midplane sections in an IBM @server BladeCenter are independent of each other (see Figure 2-3). Having a dual midplane ensures that there is no single point of failure and the blades remain operational.

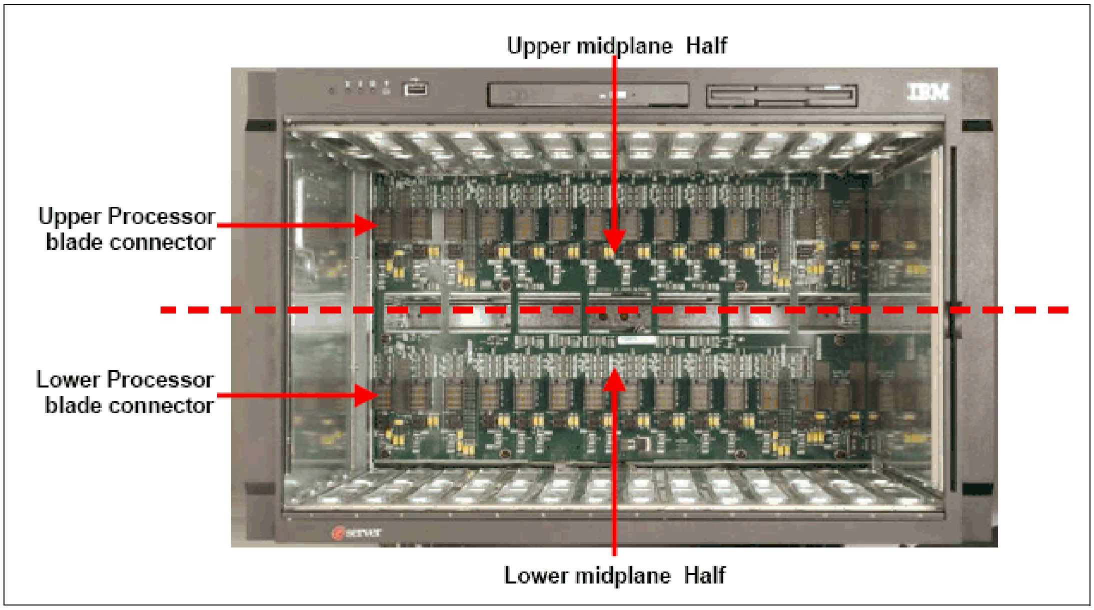

Upper midplane Half
Upper Processor blade connector
Lower Processor blade connector
Lower midplane Half

Figure 2-3 Internal picture of the upper and lower midplane of the BladeCenter chassis

2.2.2 Management Module Ethernet

Figure 2-4 illustrates the Management Module Ethernet interface. The switch modules are configured by the active Management Module through the use of a 100 Mb Ethernet interface. Each Management Module has four 100 Mb Ethernet interfaces, one for each switch module. Each switch module has two 100 Mb Ethernet interfaces, one for each Management Module.

> **Note:** On the Nortel Networks L2/3 GbESM, the management Ethernet ports on the switch are referred to as MGT1 and MGT2. For more information beyond this generic illustration, see Chapter 4, "Nortel Networks Layer 2/3 GbE Switch Module architecture" on page 21.

The following list clarifies the routing:

- Management Module 1 Ethernet 1 → Switch Module 1 Ethernet MGT1
- Management Module 1 Ethernet 2 → Switch Module 2 Ethernet MGT1
- Management Module 1 Ethernet 3 → Expansion Switch Module 3 Ethernet MGT1
- Management Module 1 Ethernet 4 → Expansion Switch Module 4 Ethernet MGT1
- Management Module 2 Ethernet 1 → Switch Module 1 Ethernet MGT2
- Management Module 2 Ethernet 2 → Switch Module 2 Ethernet MGT2
- Management Module 2 Ethernet 3 → Expansion Switch Module 3 Ethernet MGT2
- Management Module 2 Ethernet 4 → Expansion Switch Module 4 Ethernet MGT2

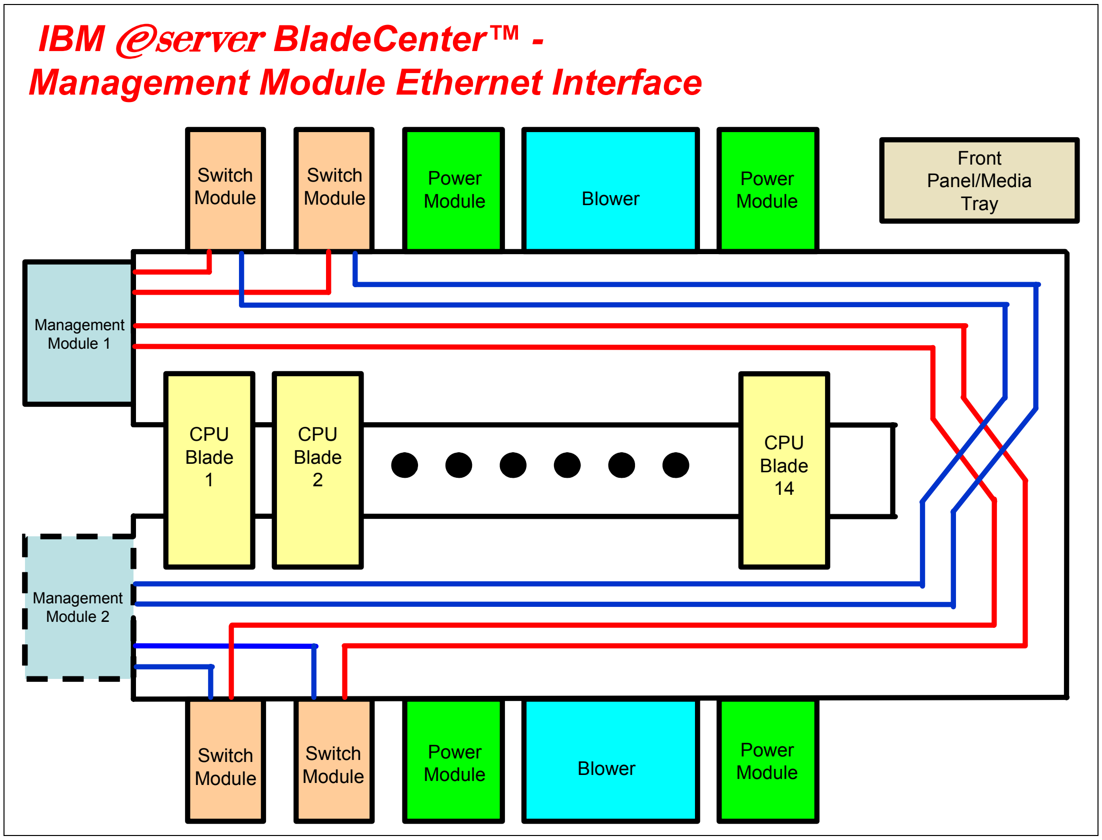


Figure 2-4 Management Module Ethernet interface

The redundant paths of the Management Module Ethernet interface are run from Management Module 2.

2.2.3 Gigabit Ethernet path

Figure 2-5 on page 10 illustrates the Gigabit Ethernet path. Each processor blade has a minimum of two and a maximum of four EtherLAN interfaces. In particular, the BladeCenter HS20 processor blade has two serializer/deserializer SERDES-based Gb Ethernet interfaces, one for each midplane connector. With a daughter card installed, two more network interfaces can be added. Each switch module (SW Module) receives one LAN input from each processor blade, for a total of 14 inputs.

> **Note:** On the Nortel Networks L2/3 GbESM, the internal Ethernet ports on the switch are referred to as MGT1 and MGT2. For more information beyond this generic illustration, see Chapter 4, "Nortel Networks Layer 2/3 GbE Switch Module architecture" on page 21.

The following partial listing illustrates the routing:

- Processor blade 1 LAN 1 → Switch Module 1 input INT1
- Processor blade 1 LAN 2 → Switch Module 2 input INT1
- Processor blade 1 LAN 3 → Expansion Switch Module 3 input INT1
- Processor blade 1 LAN 4 → Expansion Switch Module 4 input INT1
- Processor blade 2 LAN 1 → Switch Module 1 input INT2
- Processor blade 2 LAN 2 → Switch Module 2 input INT2
- Processor blade 2 LAN 3 → Expansion Switch Module 3 input INT2
- Processor blade 2 LAN 4 → Expansion Switch Module 4 input INT2

On processor blade, LAN 1 and LAN 2 are the on-board SERDES Gbit Ethernet interfaces, and are routed to Switch Module 1 and Switch Module 2, respectively, for every processor blade. LAN 3 and LAN 4 go to the Expansion Switch Modules 3 and 4, respectively, and are only to be used when a daughter card is installed. Unless a daughter card is installed in one or more processor blades, there is no need for Switch Modules 3 and 4. Further, the switch modules have to be compatible with the LAN interface generated by the processor blade. If a Fibre Channel daughter card is installed in a BladeCenter HS20 processor blade, Switch Modules 3 and 4 must also be Fibre Channel-based, and any daughter cards installed in the remaining BladeCenter HS20 processor blades must be Fibre Channel.

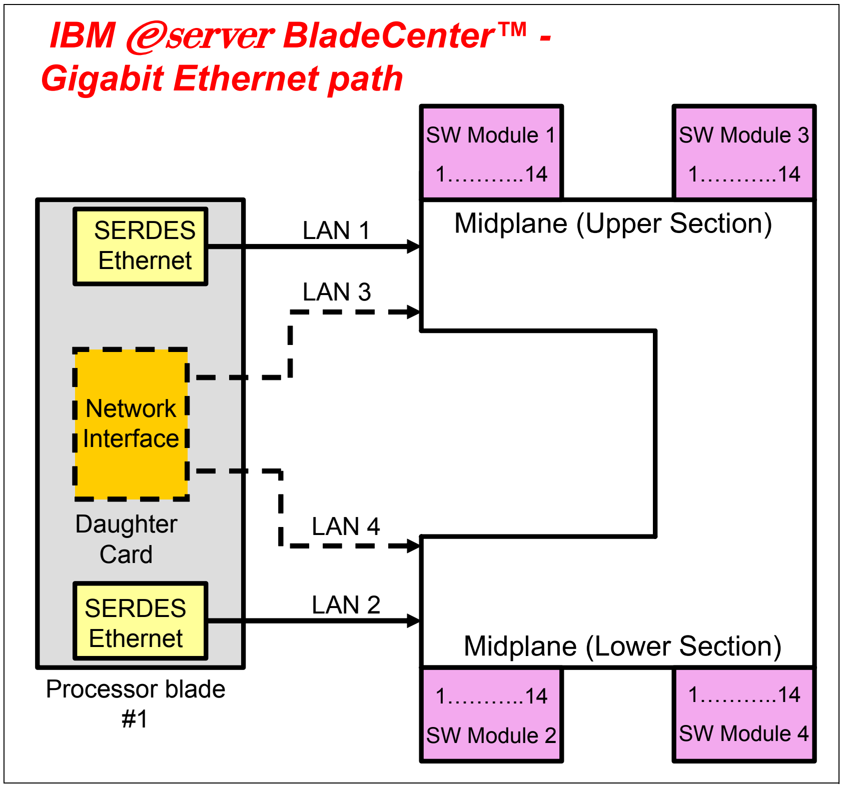


Figure 2-5 Gigabit Ethernet path

2.3 IBM @server HS20 architecture

In this section, we discuss the architectural design of the IBM @server BladeCenter HS20. This is presented as just one example of the blade design for a typical dual-processor server.

The BladeCenter HS20 uses the Intel® Lindenhurst chipset (see the HS20 architecture in Figure 2-6).

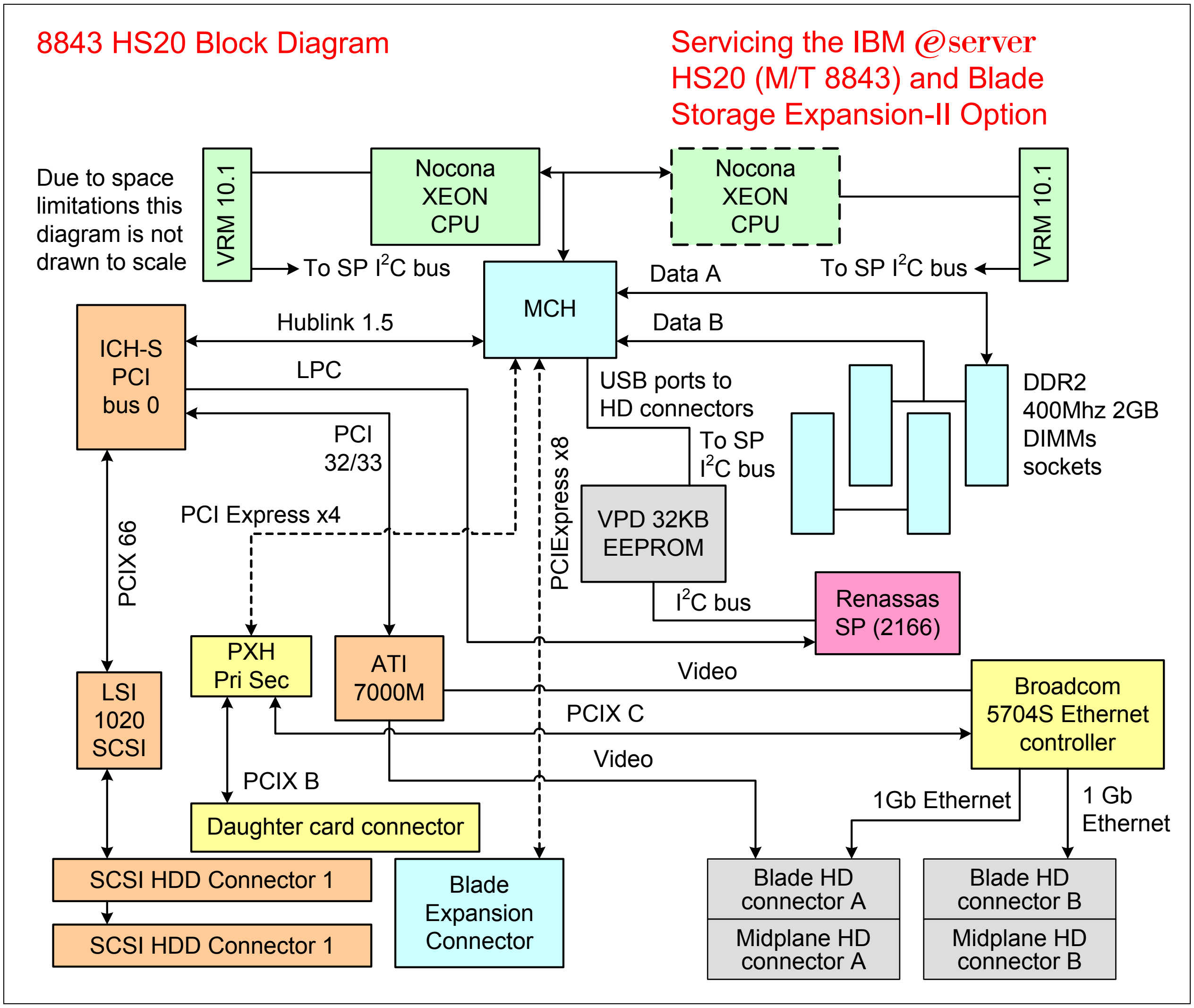


Figure 2-6 HS20 architecture

The Intel Lindenhurst chipset consists of the following components:

- Memory and I/O controller (MCH) (North Bridge)
- PXH-D
- ICH-S (South Bridge)

The Lindenhurst MCH, Memory and I/O controller provides the interface between the processors, the memory, and the PCI Express busses that interface to the other Intel chips. The Lindenhurst ICH-S (South Bridge) provides the USB interfaces, the local Service Processor interface, the POST/BIOS flash EEPROM interface, and the PCI bus interface for the ATI Radeon Mobility Video controller and LSI 1020 SCSI Host Controller. The PXH interfaces the Broadcom BCM5704S ethernet controller on its secondary bus and the daughter card on its secondary bus. I/O functions on the 8843 include Video, I2C, USB, SCSI, Gigabit Ethernet, and USB (floppy, CD-ROM (DVD), mouse, and keyboard).

The LPC bus is used to connect to the POST/BIOS EEPROM on the 8843. The size of the EEPROM is 4 MB x 8, and it contains primary BIOS, backup BIOS, and blade diagnostics.

PCI Express features include:

- PCI software compatibility
- Chip-to-chip, board-to-board implementations
- Support for end-to-end data integrity
- Advanced error reporting and handling for fault isolation and system recovery
- Low-overhead, low-latency data transfers and maximized interconnect efficiency
- High-bandwidth, low pin-count implementations for optimized performance

2.4 Stand-alone configuration tools

IBM @server BladeCenter hardware can be configured using standard software, such as a Web browser and a Telnet client, which are available on all the mainstream operating system platforms. This is possible by exploiting Web and American National Standards Institute (ANSI) interfaces that are embedded in both the management and the Ethernet Switch Modules.

A very comprehensive tool is accessible through the Web interface. This tool contains various configuration submenus, and one of them (I/O Module Tasks) lets you set up the Ethernet Switch Module. Basic settings (such as the Ethernet Switch Module IP address and the enablement of the external ports) are configured by exploiting the I2C bus. An advanced menu allows for the fine tuning of the module, by either opening another window of the Web browser or running a Java™ applet that allows for connectivity to an ANSI interface. (This requires that you have Java 2 V1.4 installed on the management system.) To achieve this, the 10/100 Mb internal link that connects the Management Module and the Ethernet Switch Modules through the BladeCenter backplane are exploited (notice that the internal network interface of the Management Module has a default static IP address of 192.168.70.126).

These more complete tools can also be accessed by pointing your Web browser, Telnet, or SSH client to the IP of the Ethernet Switch Module itself. (The default for a module that is plugged into Rear Bay 1 is 192.168.70.127. However, you can configure Dynamic Host Configuration Protocol (DHCP) based addressing.) Notice that this latter capability requires the management system to connect through the external ports (on the production LAN) of the Ethernet Switch Module and, therefore, might potentially raise concerns about security. That is why you have the capability to disable configuration control through the external ports in the I/O Module Tasks of the Management Module interface.

Figure 2-7 on page 13 illustrates the available stand-alone configuration tools.

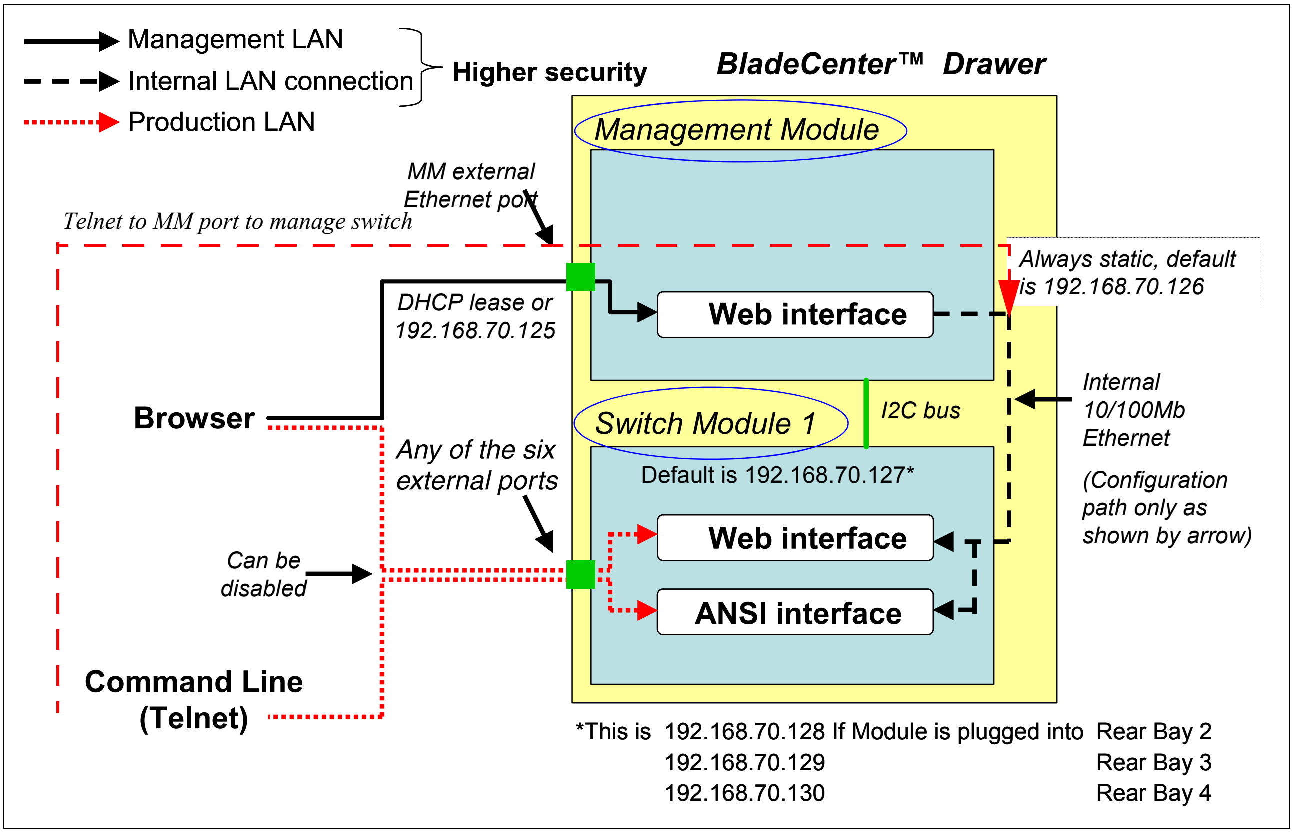

Management LAN
Internal LAN connection
Production LAN
Higher security
BladeCenter™ Drawer
Management Module
MM external Ethernet port
Telnet to MM port to manage switch
Always static, default is 192.168.70.126
DHCP lease or 192.168.70.125
Web interface
Browser
Internal 10/100Mb Ethernet
(Configuration path only as shown by arrow)
I2C bus
Switch Module 1
Default is 192.168.70.127*
Any of the six external ports
Web interface
Can be disabled
ANSI interface
Command Line (Telnet)
*This is 192.168.70.128 If Module is plugged into Rear Bay 2
192.168.70.129 Rear Bay 3
192.168.70.130 Rear Bay 4

Figure 2-7 Stand-alone configuration tools

3

Nortel Networks Layer 2/3 GbE Switch Modules

This chapter discusses the Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter and its set of features and services.

3.1 Product description

The new Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter serve as a switching and routing fabric for the BladeCenter server chassis. In addition to the Layer 2 switching capabilities, these switches introduce the expanded capabilities of Layer 3 routing. Up to four copper or fiber Gb Ethernet modules can reside in the I/O module bays of the BladeCenter chassis. The modules can be hot-plugged into an IBM @server BladeCenter without disrupting normal operations.

The Nortel Networks L2/3 GbESM connects to the server blades via the 14 internal GbE interfaces (server ports) over the BladeCenter midplane. It supplies six external copper or multimode fiber GbE interfaces for outside communication (shown in Figure 3-1). The switch is managed via two internal 100 Mbps ports for communication to the BladeCenter management module. A RS232 serial console management interface is also available.

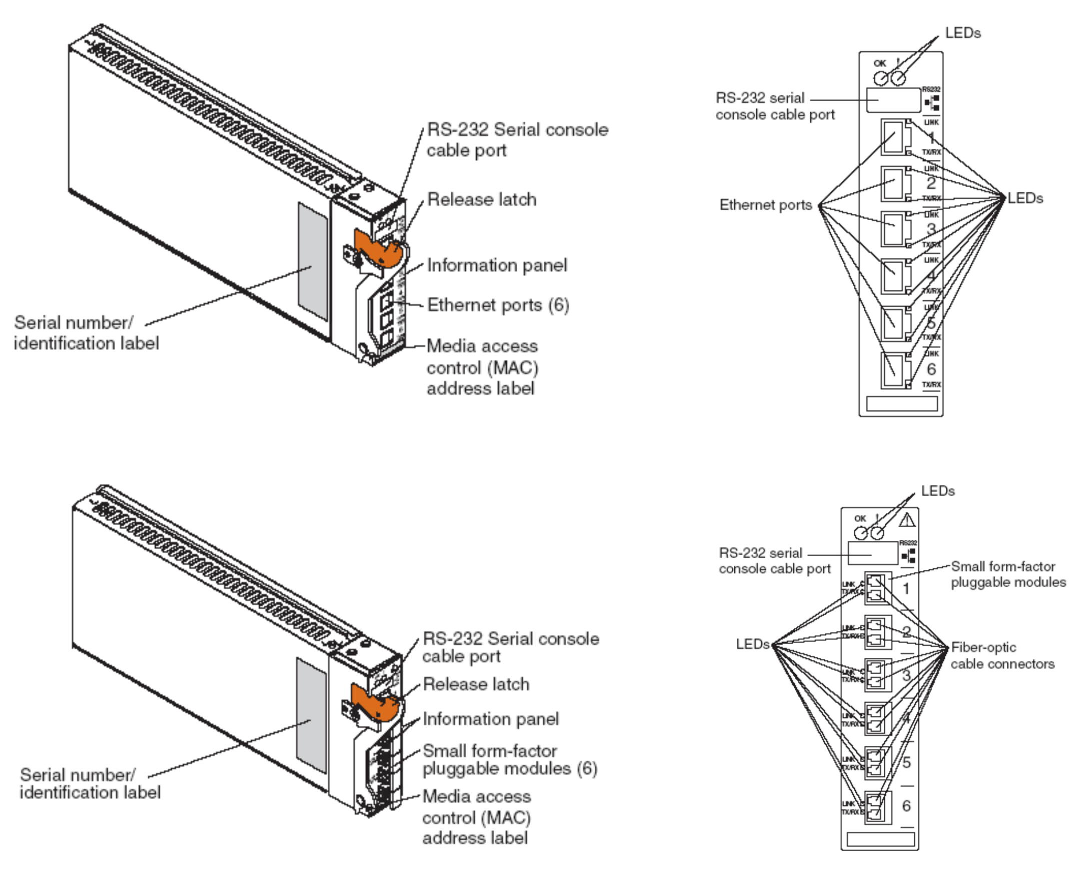

RS-232 Serial console cable port
Release latch
Information panel
Ethernet ports (6)
Serial number/ identification label
Media access control (MAC) address label
LEDs
OK
RS-232 serial console cable port
RS232
Ethernet ports
LINK
TX/RX
1
2
3
4
5
6
LEDs
RS-232 Serial console cable port
Release latch
Information panel
Small form-factor pluggable modules (6)
Serial number/ identification label
Media access control (MAC) address label
LEDs
OK
RS-232 serial console cable port
Small form-factor pluggable modules
LEDs
Fiber-optic cable connectors

Figure 3-1 Nortel Networks L2/3 GbESM connections

Full Layer 2 switching and Layer 3 routing provide flexible in-chassis traffic management and security. The Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter provides full Layer 2 switching with availability capabilities such as advanced spanning tree protocols, Link Aggregation Control, Cisco Etherchannel, and 802.1Q VLANs, application delivery and performance features such as granular QOS (Differentiated Service Code Point 802.1p), Internet Group Management Protocol (IGMP) snooping, and multicast.

In particular, the switch modules support up to 16,384 MAC addresses, 4,096 address resolution protocol (ARP) entries, and up to 2,048 dynamic route entries to ensure a high level of support for a number of users. The IEEE 802.1D Spanning Tree Protocol (STP) support can be enabled or disabled on a per-port basis. Multiple instances of STP are supported (that is, 16 STP groups). Virtual Local Area Network (VLAN) support includes 802.1Q tagged VLANs and support for IEEE 802.3 support on six external ports for up to three static trunk groups. Dynamic trunking using LACP as well as static trunking is supported.

Adding full Layer 3 routing to the integrated switch module adds more power, flexibility, and security capabilities to the IBM @server BladeCenter. With the integrated switch module in the BladeCenter, network traffic can be managed much more efficiently. Broadcast traffic can be contained in the blade server by placing the 14 blade servers on different subnets while allowing communication between each without using the bandwidth of the external ports to send traffic to and from an external Layer 3 device.

Security features provide added protection for switch configuration data, while packet filtering helps secure and segment sensitive traffic or network access. Support for Simple Network Management Protocol (SNMPv3), Secure Shell (SSHv2), and Hypertext Transfer Protocol over Secure Socket Layer (HTTPS) supply protection for sensitive switch configuration data. Multilevel access and defined access policies help secure the switch against unauthorized management access. Support for Remote Authentication and Remote Authentication Dial-in User Service Protocol (RADIUS), and Terminal Access Controller Access Control System (TACACS+) gives enterprises the freedom to use current security databases.

Layer 3 filtering (IP and application type) at line rate in the chassis enhances security and simplifies provisioning. The risk of traffic finding a route to a denied destination is reduced if Layer 3 routing is contained in the switch module in the chassis. Without Layer 3 filtering, several external switches might need configuration to filter traffic to limit access between one server blade and another if the traffic flows through upstream devices.

The following routing standards are supported:

- Routing Information Protocol version 1 (RIPv1), and version 2 (RIPv2)
- Border Gateway Protocol version 4 (BGPv4)
- Open Shortest Path First version 2 (OSPFv2)

Maximum bandwidth and network flexibility with uplink support for six Gigabit Ethernet ports/switch (fiber or copper). The Nortel Networks L2/3 GbESM is designed to be able to route, filter, and queue traffic so that no data is lost, dropped, or delayed. Applications get the bandwidth they need, when they need it, with little or no delay or jitter.

Even with all the value that an integrated switch module can provide, performance for a BladeCenter can be limited if the switch module cannot provide adequate ingress and egress bandwidth. The Nortel Networks L2/3 GbESM is the only BladeCenter switch module that offers six Gigabit Ethernet uplink ports for maximum throughput, supporting full Layer 2 through Layer 3 wire-speed packet forwarding for all connections. In addition, the flexibility of both copper and fiber ports allows for optimized use in hybrid installations or for situations where one switching infrastructure is more economical than another.

Unmatched High Availability support and field proven Resiliency High availability support is built in at both Layer 2 and Layer 3 in the Nortel Networks L2/3 GbESM to reduce single points of failure when it comes to enabling reliability and performance of the network.

At Layer 2 Link Aggregation Control (802.3), Rapid Spanning Tree, Fast Uplink Convergence, Port Fast Forwarding, 802.1Q VLANs, Broadcast Storm Control, and Native Link Failover with NIC teaming are supported.

At Layer 3, special configurations of Virtual Router Redundancy Protocol (VRRP) allow all switches in the VRRP group to concurrently process traffic by using multiple instances of VRRP. Such configurations enable maximum switch performance while also ensuring seamless failover in the unlikely event of a failure. VRRP Hot Standby is also supported to enable effective use of NIC Teaming in Layer 3 network topologies much as Trunk Failover facilitates HA designs with NIC Teaming at Layer 2.

3.2 Value proposition

This section discusses the value of using the Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter for your IBM @server BladeCenter.

Product strength

The product provides strengths such as:

- Provides full interoperability into existing Nortel and Cisco data centers with the BladeCenter integrated GbE switch module.
- Integrates Nortel networking capabilities to reduce data center complexity and increases networking manageability and availability.
- Leverages the leadership capabilities BladeCenter Alliance Partners to provide the most technological choices.

Leadership features and function

The leadership features and function include:

- IBM @server BladeCenter delivers with the Nortel GbESM, full Layer 2 switching and Layer 3 switching (routing) functionality as well as Layer 4 filtering and related services.
- The switch module runs Alteon Operating System and appears as any other product from Nortel's Alteon product line to the data center's network management tools. In addition, Nortel is pursuing a unified command line syntax across its data products, known as the NNCLI (Nortel Networks CLI), which will be available on the L2-3 switch late in 2005.

Competitive advantage

The product delivers a competitive advantage by delivering:

- Full integration of Ethernet switching, reducing infrastructure complexity
- Six external copper or fiber option
- Upgrade path to full Layer 4-7 services
- Price leadership

3.3 Supported hardware

Table 3-1 lists the following IBM hardware platforms which support Nortel Networks Layer 2/3 Copper Gigabit Ethernet Switch Module for IBM @server BladeCenter (26K6530) and Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter (26K6531).

Table 3-1 Supported platforms

System name	Machine type	Model
BladeCenter	8677	All
BladeCenter	7967	All
BladeCenter T	8720	All
BladeCenter T	8730	All

Product shipment group

The items that ship with either switch module are:

- Nortel Networks Layer 2/3 Copper Gigabit Ethernet Switch Module for IBM @server BladeCenter (26K6526) or Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter (26K6529)
- Serial Console Cable (FRU 02R9365)
- Installation publication, including Documentation CD
- Safety flyer
- Software License Agreement
- Six small form-factors (SFPs) are pre-installed into the Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter (26R0808)

4

Nortel Networks Layer 2/3 GbE Switch Module architecture

This chapter provides a system overview of the Nortel Networks Layer 2/3 GbE Switch Module.

4.1 Nortel GbESM architecture overview

The Nortel Networks Layer 2/3 GbE Switch Module is a fully functional Layer 2 and 3 switch that includes Layer 4 awareness and capability. Figure 4-1 shows the architecture overview of the Nortel Networks Layer 2/3 GbE Switch Module.

The Nortel GbESM has 14 internal 1 Gbps links to connect to blade servers and six external Gigabit ports to connect to upstream switches. The switch module has two 100 Mbps connections to the Management Modules. You can manage the Nortel GbESM through the connection between the Nortel GbESM and the Management Module. You can also manage the Nortel GbESM like other switches with the RS232 console port that looks similar to a USB port. The console port is a service port to which you can connect a terminal or PC in order to configure the software through the command-line interface (CLI) or to troubleshoot problems with the switch.

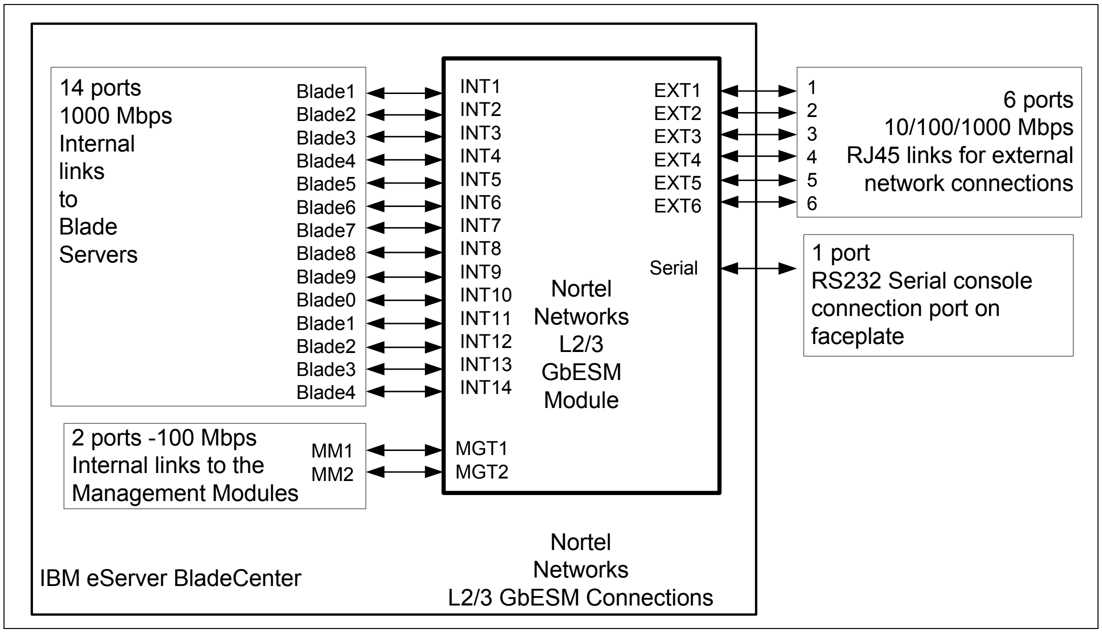


Figure 4-1 Nortel Networks L2/3 GbESM architecture overview

Figure 4-2 on page 23 shows the architecture for Ethernet connectivity. The two Nortel GbESMs can be housed within the BladeCenter chassis. Each Nortel GbESM provides six uplink ports, which can be grouped to support 802.3ad Link Aggregation. The blade server has two NICs, with NIC 1 connecting to Nortel GbESM 1 and NIC 2 connecting to Nortel GbESM 2. The links connecting the blade servers to the Nortel GbESMs are on the backplane of the BladeCenter chassis. The Nortel GbESM has two links to the Management Modules. Each link connects to a different Management Module.

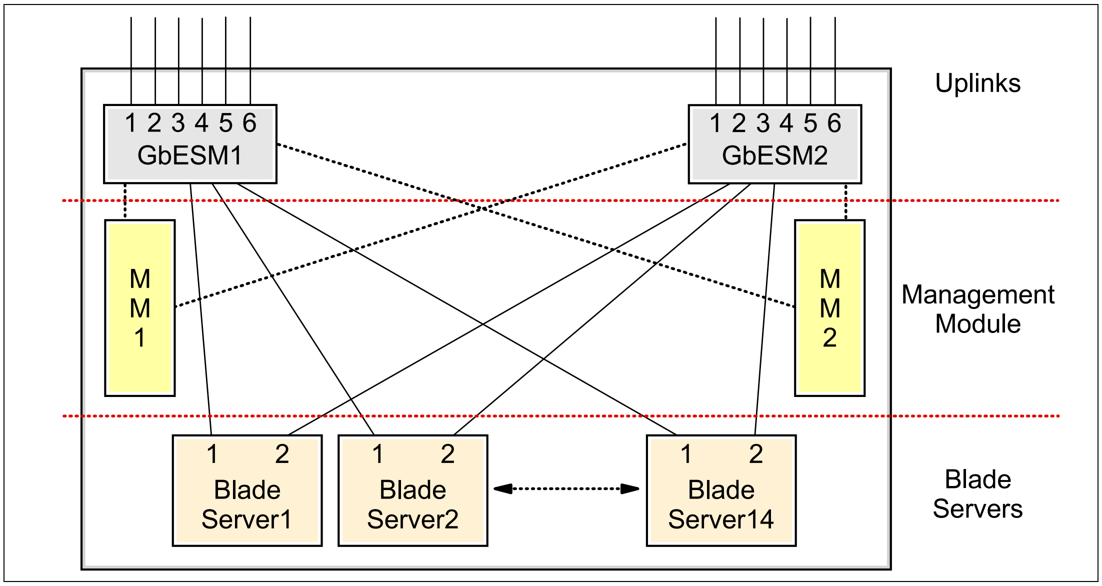


Figure 4-2 BladeCenter Ethernet connectivity

Internal Layer 2 traffic flow in the Nortel Networks L2/3 GbESM

Figure 4-3 shows the internal Layer 2 traffic flow in the Nortel GbESM. The hard coded filter in the Nortel GbESM blocks all traffic between the external ports and the Management Module ports. Two Nortel GbESMs in the same BladeCenter chassis exchange Layer 2 frames across the Management Module. The Nortel GbESM processes BPDUs that reach it via the Management Module if Spanning Tree is enabled for the Management Module ports. This is rarely necessary.

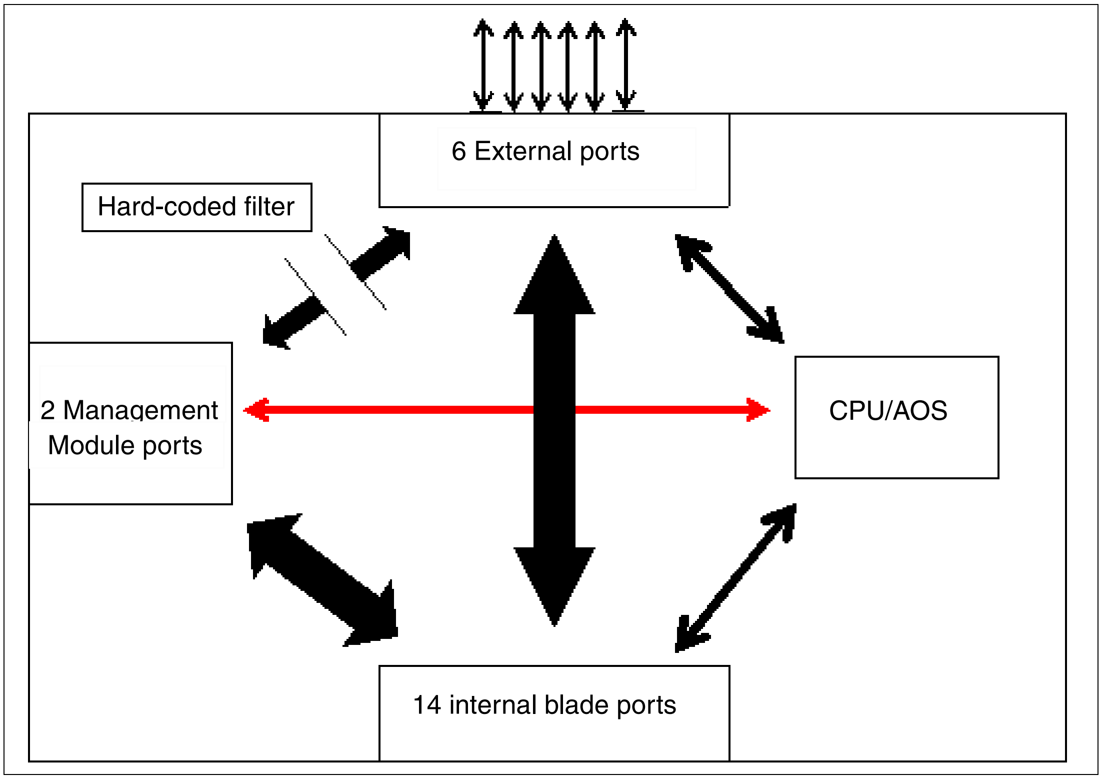


Figure 4-3 Layer 2 frames flow in the Nortel Networks L2/3 GbESM

Figure 4-3 on page 23 also indicates the following:

- Two Nortel GbESMs in the same BladeCenter chassis can ping or telnet to each other without connecting external ports. They cannot pass user data to each other via this path, which passes traffic through the Management Module.
- The internal blade ports cannot be on the same VLAN as the Management Module ports. As a result, the blade servers on the production network must be on a different IP subnet than the Management Module and other devices which are on the management network.

4.2 Nortel Networks L2/3 GbESM block diagram

Figure 4-4 shows the block diagram of the Nortel Networks Layer 2/3 GbE Switch Module.

The Nortel GbESM has two Gigabit Ethernet Aggregator (GEAs) for switching. It has 1 MB on chip cache for packet buffers and supports 20 Gigabit Ethernet ports (14 internal ports and six external ports). The two GEAs are interconnected with 10 Gigabit proprietary link, which is shown as the *10G HiGig link* in Figure 4-4. HiGig is a proprietary protocol from IBM.

GEA0 supports eight Gigabit Ethernet ports (two internal connections-5421s to the Management Modules and six external ports). GEA1 supports the remaining 12 internal ports. The connection between the 5421s and the Management Module links up at 100 Mbps.

Copper ports use six external 1000BASE-T RJ-45 connectors. Fiber ports use six 1000BASE SX SFP transceivers that are included with the GbE switch module.

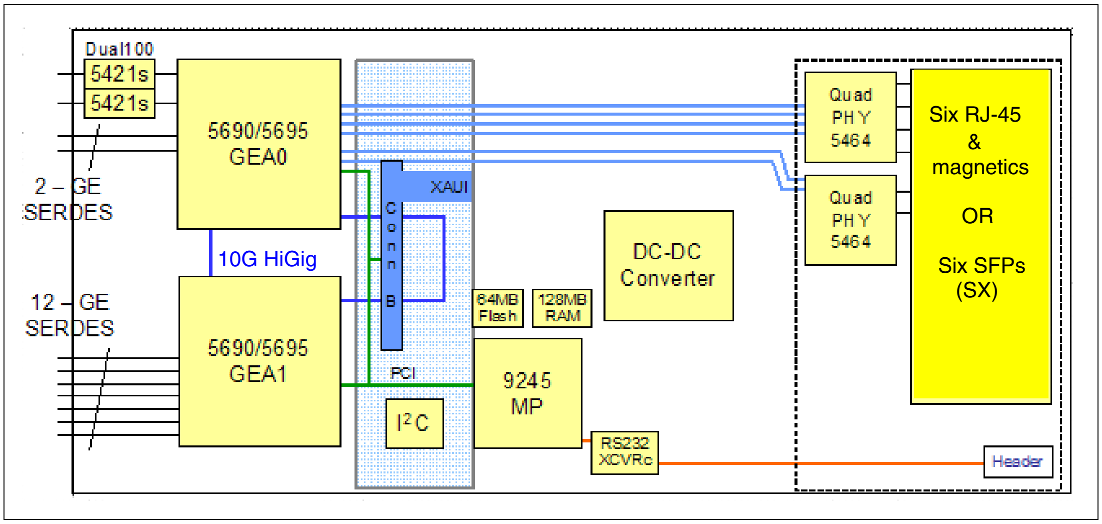


Figure 4-4 Nortel Networks L2/3 GbESM block diagram

4.2.1 Nortel Networks L2/3 GbESM ports specific roles

Figure 4-5, Figure 4-6, and Figure 4-7 on page 26 show different examples of the port connections to the Nortel GbESM(s) within the IBM @server BladeCenter. We then discuss the specific roles and restrictions for the various ports.

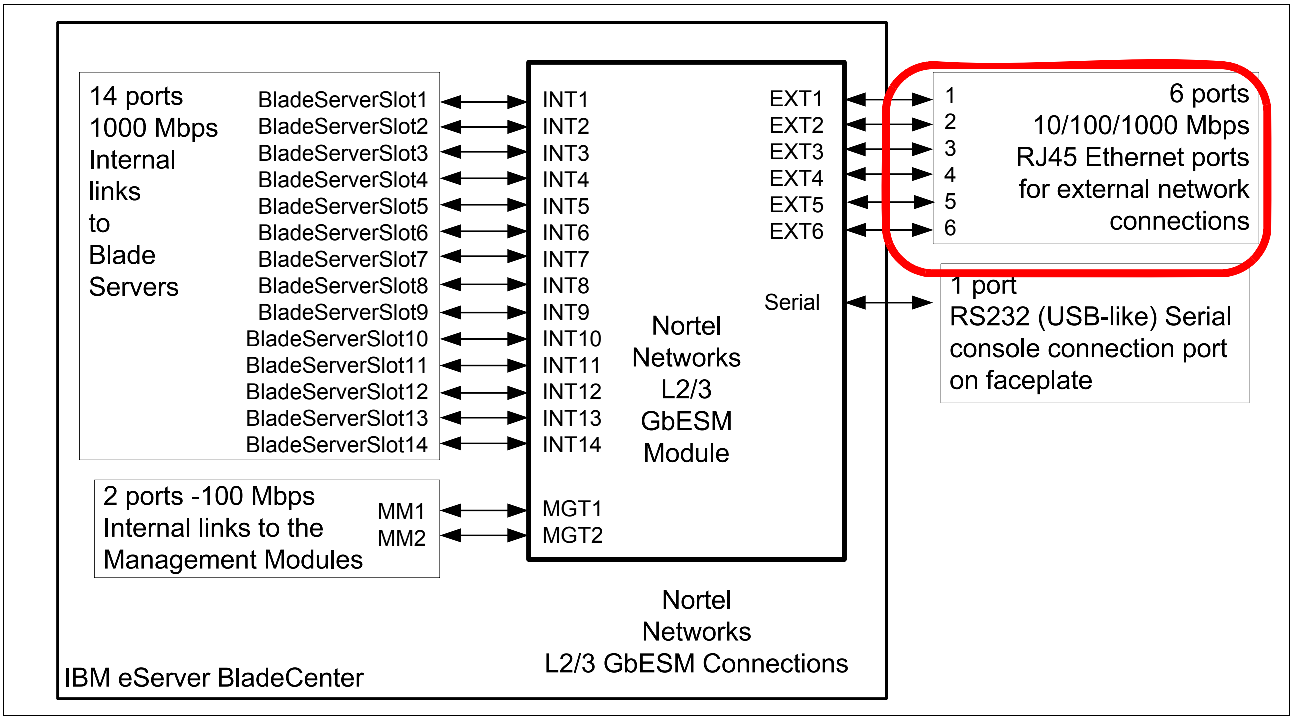


Figure 4-5 Connections on the Nortel GbESM

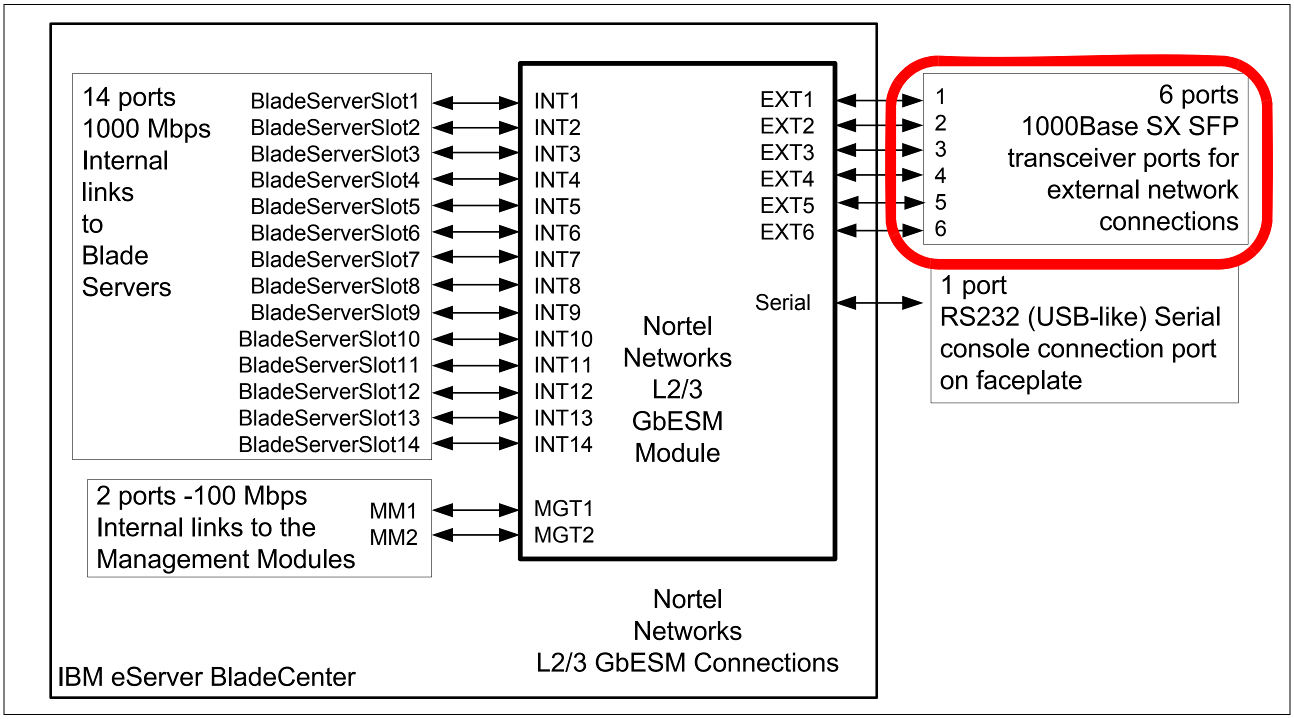


Figure 4-6 Connections on the Nortel GbESM

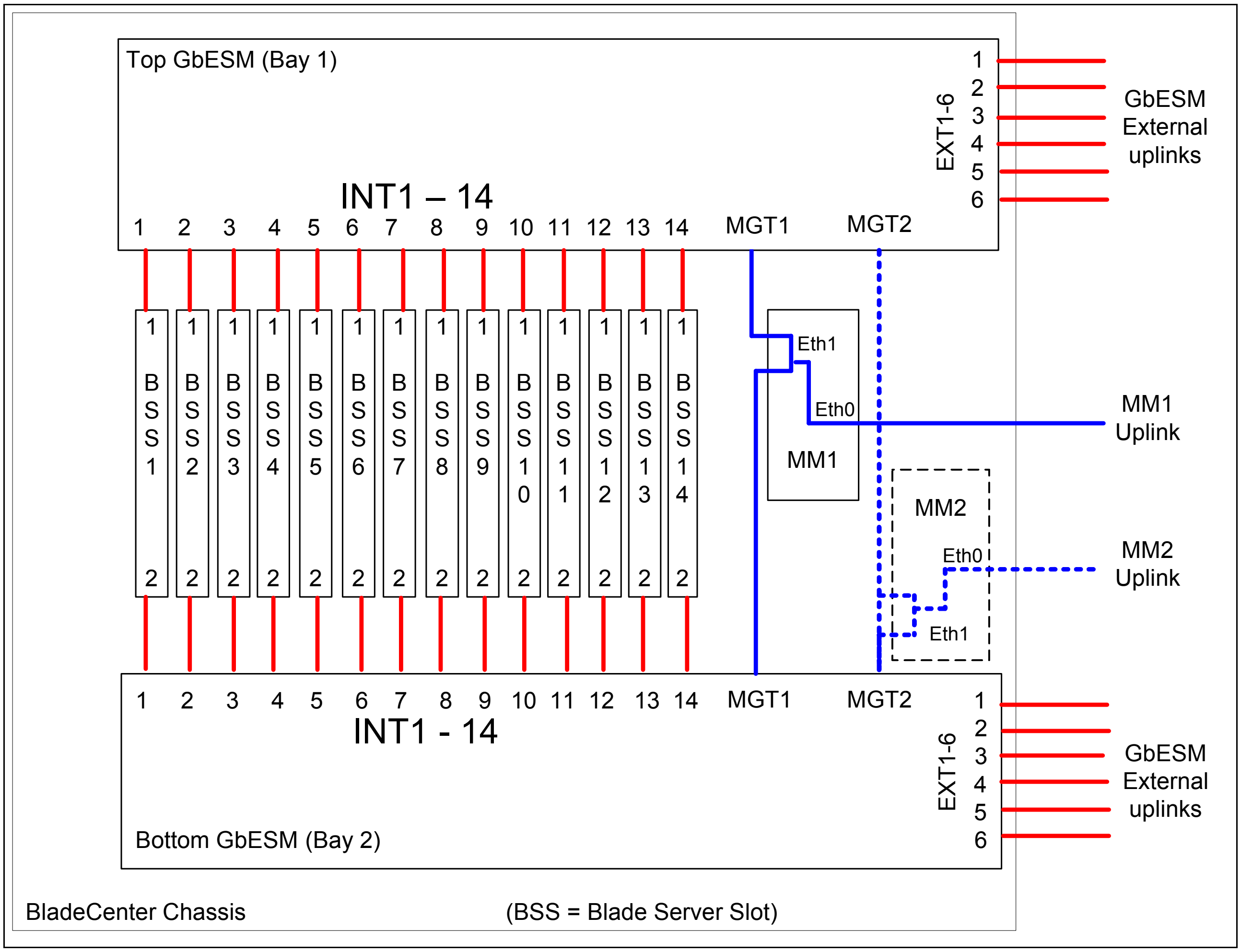


Figure 4-7 Overall view of port connections within an IBM @server BladeCenter

Ports INT1 through INT14: Connects to blade server slots 1 through 14, respectively:

- Preset default values for ports going to the blade servers (includes ports INT1 through INT14):
  - IEEE 802.1Q tagging is enabled
  - Default VLAN is VLAN1
  - VLAN 4095 is reserved for Serial over LAN
- Hard-coded to Auto negotiation, but only support 1000/full duplex to the blade servers. This cannot be changed at this time, but future revisions of code may support the ability to set these ports to a no negotiate condition and force the link to 1000/full.

- Spanning Tree (STP) is disabled by default for all internal ports.

Preset default values for ports going to the Management Modules (includes ports MGT1 and MGT2):

- Speed is hard-coded at 100 full and cannot be changed.
- Ports MGT1 and MGT2 cannot be disabled.
  - This is by design to ensure that the links to the BladeCenter Management Modules are not inadvertently or intentionally brought down by the administrator.
  - Note that only one of these ports (MGT1 or MGT2) is active at one time (only one Management Module is active at any given time).
- Both ports are hard-coded as untagged VLAN 4095 (internal management VLAN).
- Nortel has implemented a hidden filter (not visible or controllable by the administrator) that prevents any packet entering one of the uplink ports or the internal ports (INT1 -14 and EXT1 - 6) from exiting toward the Management Module ports (MGT1 - 2) and vice-versa. This filter is implemented to ensure isolation of the internal BladeCenter management network.

Ports EXT1 through EXT6: Connects to external ports 1 through 6, respectively:

- Preset default values for ports going to external connections (includes ports EXT1 through EXT6): Untagged and configured on VLAN 1.
- These ports default to Disabled when in a new IBM @server BladeCenter. You must use the Management Module Web interface, under I/O Module tasks Advanced settings, to set External Ports to Enabled to bring them up the first time.

RS232 Console port:

- Default settings:
  - Baud rate: 9600
  - Data bits: 8
  - Parity: None
  - Stop bits: 1
  - Flow control: None
  - Emulate: VT100
- Serial Console Cable - (FRU 02R9365).

  This USB-style connector enables connection to the GbE switch module.

The management VLAN IP address information is not lost during factory reset

The management VLAN IP address information is not lost as long as Preserve new IP configuration on all resets is enabled on the Management Module.

As a direct result of a feature being enabled on the Management Module (under I/O Modules Advanced Setup), after a Nortel GbESM is cleared (reload or through the GUI), the BladeCenter Management Module provides its currently saved IP information for that Nortel GbESM. This is to help ensure that the Nortel GbESM can always be accessed over from the Management Modules. This action (providing or not providing the Nortel GbESM its default address) can be partially controlled from the Management Modules Web interface.

See“Enabling Nortel Networks L2/3 GbESM uplink ports through the Management Module” on page 48 for details about enabling or disabling the feature called Preserve new IP configuration on all resets.

Also, if you change this setting to disabled, it is assumed that you plan on managing the Nortel Networks L2/3 GbESM via its own uplinks.

The default Nortel GbESM IP addressing that is provided by the Management Module for a new IBM @server BladeCenter is as follows:

- Switch bay 1: 192.168.70.127/24
- Switch bay 2: 192.168.70.128/24
- Switch bay 3: 192.168.70.129/24
- Switch bay 4: 192.168.70.130/24

Based on certain interactions within the IBM @server BladeCenter, it is usually *not* recommended to change the management IP address directly on the Nortel GbESM, but instead, only change it through the Management Module Web-based GUI.

5

Nortel Networks L2/3 GbESM management and administration

In this chapter, we discuss tools, techniques, and applications that help with the management and deployment of the Nortel GbESM in an IBM @server BladeCenter. We also discuss the management paths and rules for connecting to and accessing the Nortel GbESM.

> **Note:** As noted elsewhere in this document, the information herein applies to the 6-port Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter.

5.1 Nortel Networks L2/3 GbESM management connectivity

In this section, we look at the basic management connectivity and management pathways to the Nortel GbESM., as shown in Figure 5-1.

Important: Properly managing the Nortel GbESM in the IBM @server BladeCenter actually requires proper management of the Management Module within the BladeCenter chassis. In other words, it is virtually impossible to successfully deploy the Nortel GbESM if you do not understand and properly configure certain settings in the Management Module, as well as the necessary Nortel GbESM configurations.

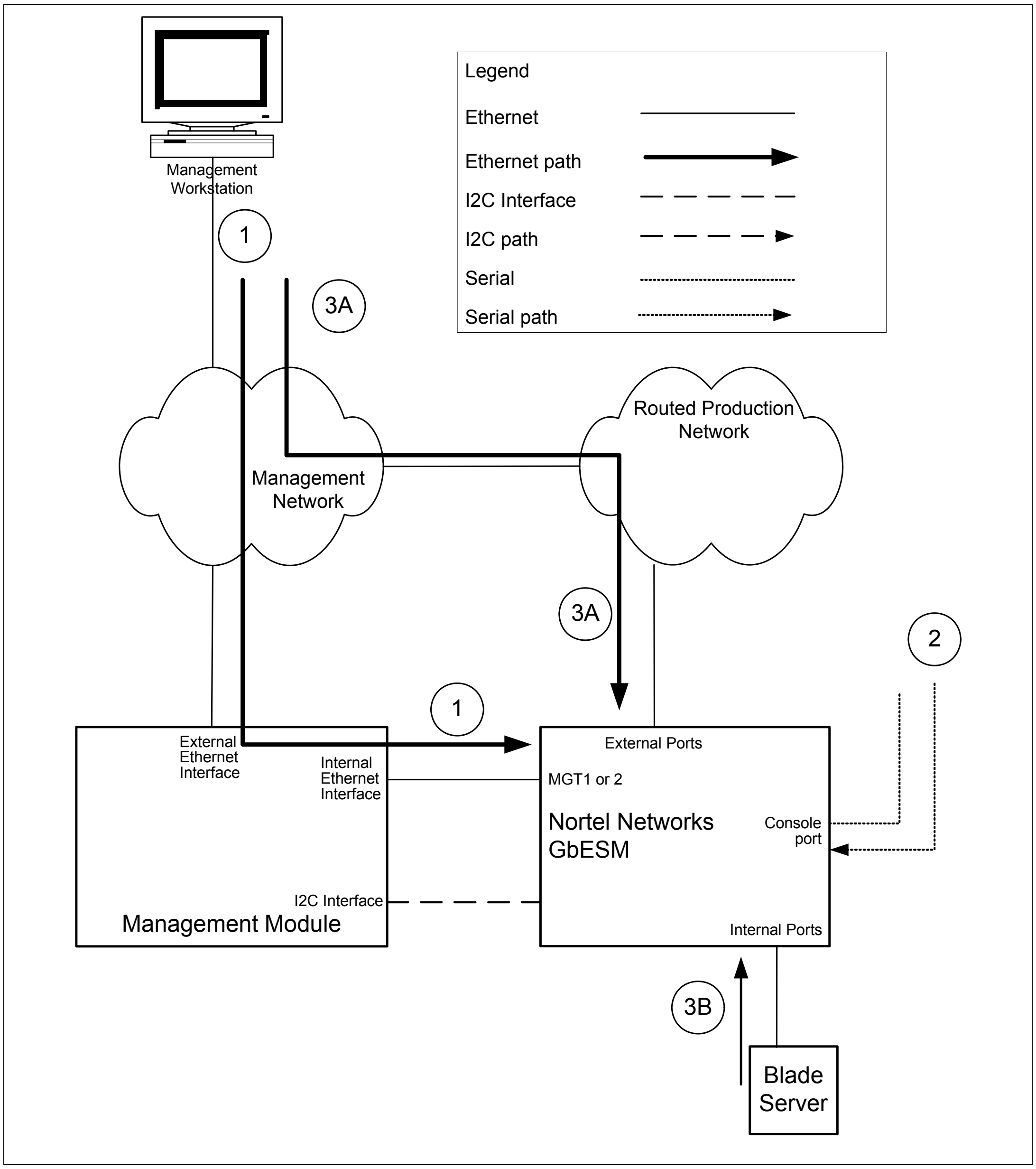


Figure 5-1 Management paths to the Nortel Networks L2/3 GbESM

5.1.1 Out-of-band management

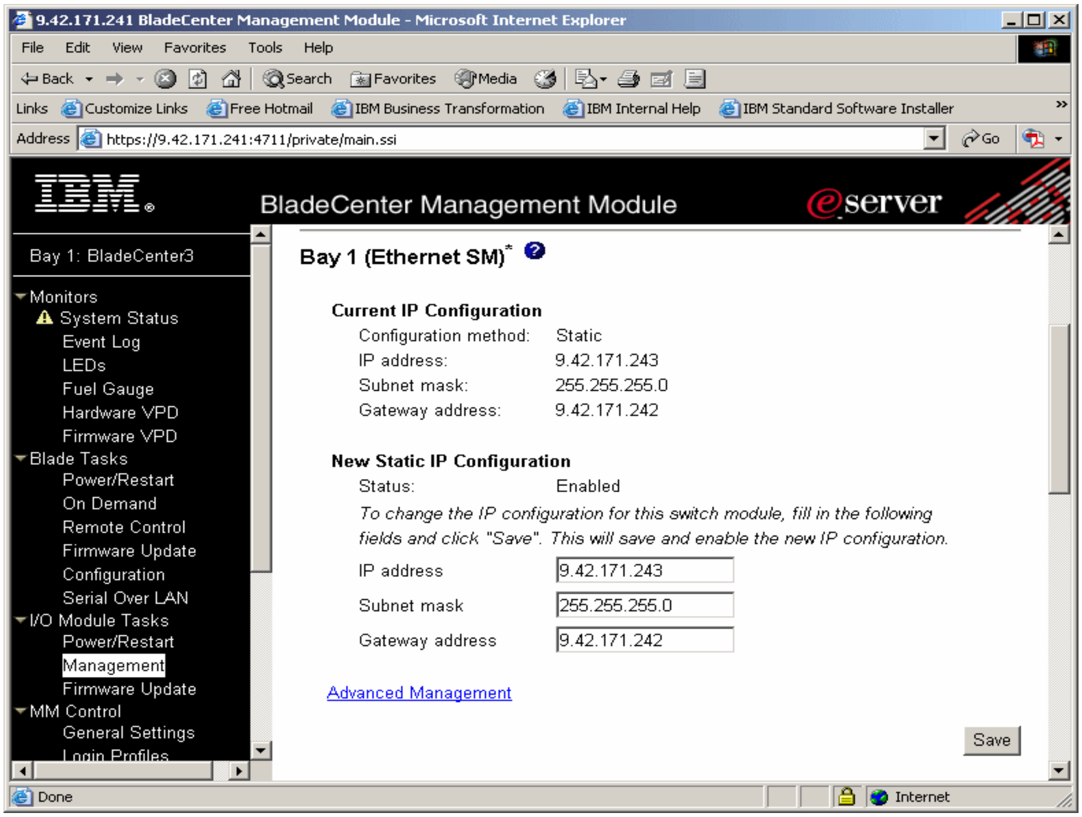
It is common to provide a (physically) separate management interface for all of the devices and to carry only management traffic. This is referred to as *out-of-band management* and is sometimes a separate Ethernet connection (path 1) or a whole different physical connection such as the console port (path 2).

Management Module (Path 1)

The IBM @server BladeCenter comes with at least one Management Module. The Management Module supports an external Ethernet interface, which is used to manage the Blade servers, Ethernet switches, and the Management Module itself. Within the IBM @server BladeCenter, management traffic flows through a different bus, the I2C bus, as shown in the Figure 5-1 on page 30.

On the Nortel GbESM, the Ethernet management (MGT1 and MGT2) ports which connect the switch to the Management Module are placed in VLAN 4095. It is not possible to change this. It is also not possible to reach VLAN 4095 from any of the other internal or external ports on the switch. This is a deliberate design constraint. It is intended to enforce isolation of the Management Module network (VLAN) from any other networks (VLANs) that are configured on the switch. This implies that the Blade servers should not be on the same VLAN nor the same IP subnet as the Management Module. Placing the servers on the same subnet as the Management Module can have unexpected and undesirable results.

The first step in configuring the Nortel GbESM is to assign the IP address of the MGT ports through the Web interface of the Management Module (Figure 5-2).


9.42.171.241 BladeCenter Management Module - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Links Customize Links Free Hotmail IBM Business Transformation IBM Internal Help IBM Standard Software Installer
Address https://9.42.171.241:4711/private/main.ssi
Go
IBM
BladeCenter Management Module
@server
Bay 1: BladeCenter3
Monitors
System Status
Event Log
LEDs
Fuel Gauge
Hardware VPD
Firmware VPD
Blade Tasks
Power/Restart
On Demand
Remote Control
Firmware Update
Configuration
Serial Over LAN
I/O Module Tasks
Power/Restart
Management
Firmware Update
MM Control
General Settings
Login Profiles
Done
Internet

Bay 1 (Ethernet SM)*

Current IP Configuration

Configuration method: Static
IP address: 9.42.171.243
Subnet mask: 255.255.255.0
Gateway address: 9.42.171.242

New Static IP Configuration

Status: Enabled

To change the IP configuration for this switch module, fill in the following fields and click "Save". This will save and enable the new IP configuration.

IP address: 9.42.171.243
Subnet mask: 255.255.255.0
Gateway address: 9.42.171.242

Advanced Management

Save

Figure 5-2 Configuring the Nortel MGT port IP address using the Management Module Web interface

Further configuration of the Nortel Networks L2/3 GbESM module is performed by using Telnet (for the Command Line Interface) or a Web browser (for the Browser Based Interface) and accessing the address of the MGT1 or 2 ports.

Note: It is recommended (and easier) to use a server or mobile computer that is external to the IBM @server BladeCenter chassis to perform initial configuration of the Nortel Networks L2/3 GbESM module. The server or mobile computer should be able to open the Web interface of the Management Module. It then can reach the switch when the switch has an appropriate IP address configured. This address must be within the same subnet as both the internal and external IP addresses of the Management Module.

Serial port (Path 2)

The Serial port is used for out of band management of the switch. It is useful to allow access to the CLI when all other paths are not working. It is possible to connect the serial port to a terminal server if desired; this allows out-of-band access to be easily provided to multiple devices.

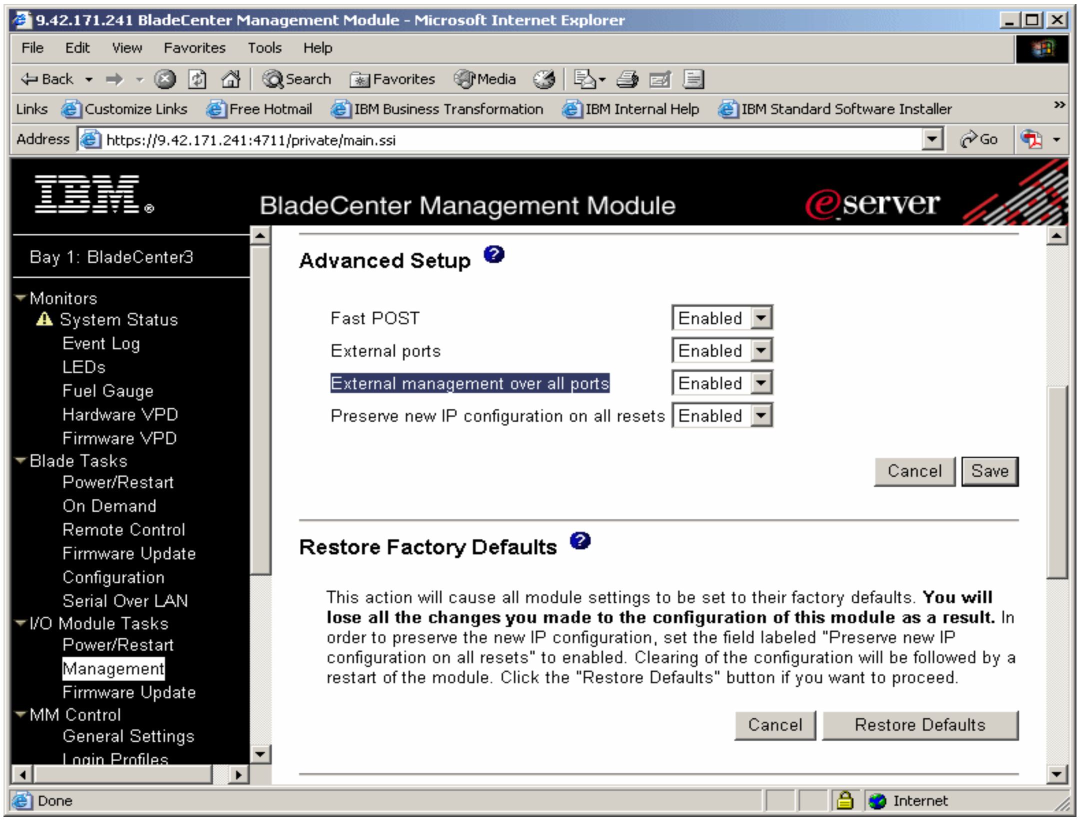
The console cable that is required to use this port is included with the switch when it is shipped. The cable has a RS232 USB-form plug on one end and a DB-9 plug on the other end. The DB-9 is intended to be attached to a standard serial port such as on a mobile computer or modem. Standard terminal emulation software should be used with these settings: 9600 baud; no parity; 8 data bits; 1 stop bit (9600,N,8,1).

5.1.2 In-band management

The second mode of operation that is commonly used is *in-band management*. In this case, the management traffic passes through the data traffic path (the Nortel Networks L2/3 GbESM EXTernal and INTernal ports).

External Ethernet ports (Path 3A)

The external ports can be used to provide management access to the switch from outside the IBM @server BladeCenter chassis. In order to use this path, the “External management over all ports” item in the Management Module configuration must be enabled (Figure 5-3 on page 33).



Advanced Setup

Fast POST	Enabled
External ports	Enabled
External management over all ports	Enabled
Preserve new IP configuration on all resets	Enabled

Restore Factory Defaults

This action will cause all module settings to be set to their factory defaults. **You will lose all the changes you made to the configuration of this module as a result.** In order to preserve the new IP configuration, set the field labeled "Preserve new IP configuration on all resets" to enabled. Clearing of the configuration will be followed by a restart of the module. Click the "Restore Defaults" button if you want to proceed.

Figure 5-3 Enabling management over all ports using the Management Module Web interface

Internal Ethernet ports (Path 3B)

The internal ports can be used to provide management access to the switch from the server blades in the same chassis.

In-band management considerations

In order to use in-band management paths, you must configure at least one additional IP address on the Nortel Networks L2/3 GbESM beyond the address that is provided through the Management Module and attached to VLAN 4095. This additional IP address should be attached to one of the active VLANs configured on the switch and is discussed in detail in 7.5.3, “Base configuration common to all examples” on page 68.

Using the `mnet` command on the Ethernet switch, it is possible to limit management access to the switch to management stations within a defined range of IP addresses.

> **Note:** This command limits all IP-based management access regardless of which path is involved. Thus, you should use it with care. It is possible to lock out access via the management module (MGT ports) using this command.

5.2 Nortel Networks L2/3 GbESM user interface

This section discusses the management interface of the switch module and what each task represents. To configure and manage the switch module, you can use the following interfaces:

- IBM @server BladeCenter Management Module and I2C

  Management functions that are necessary for initial setup are provided through the Management Module Web interface. I2C is the communication that is used between the Management Module and Ethernet switch.

- Command-line interface (CLI)

  You can configure and monitor the switch from the CLI, which is accessible through Telnet or SSH from a remote management station. You can also access the CLI through terminal emulation software on a management station directly connected to the switch module console port.

- Browser Based Interface (BBI)

  You can use the Browser Based Interface to manage and monitor the switch using a standard Web browser via HTTP. It provides a graphical means of viewing and configuring the switch's characteristics.

5.2.1 IBM @server BladeCenter Management Module and I2C

The Management Module Web interface is the only mechanism for performing certain management functions, including:

- Configuring the management IP address of the switch
- Enabling or disabling the external ports and management via these ports
- Configuring Power On Self Test (POST) options
- Remotely turning power to the switch on or off

All of these functions use the I2C interface when they need to communicate with the switch module. The use of the Management Module to configure Ethernet switches is documented in detail in the *Nortel Networks Layer 2/3 GbE Switch Module Installation Guide*.

5.2.2 Command-line interface

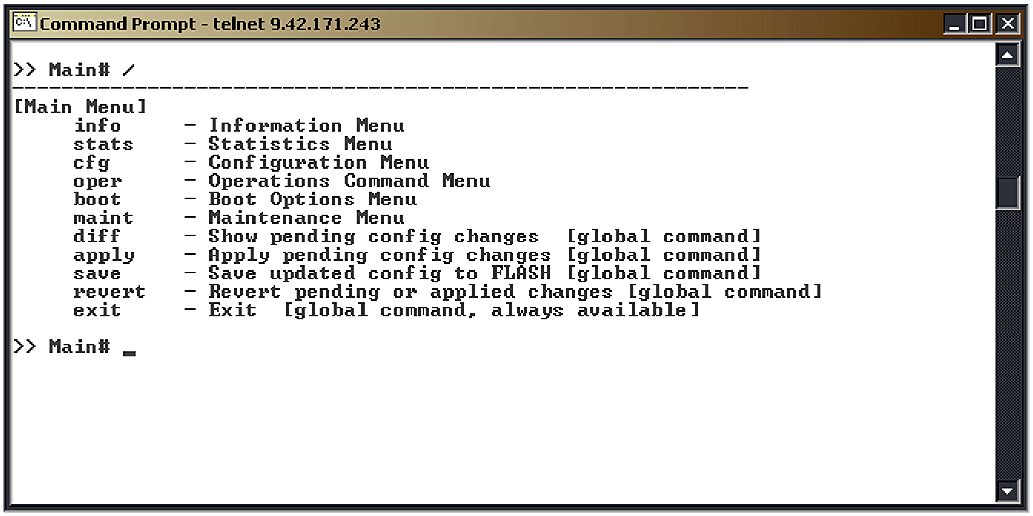
The command-line interface CLI is more flexible for configuring the switch than the BBI. It is scriptable, requires less overhead to run, and because it is a Telnet session, it can be run from any operating system (whether or not it is graphical).

Main Menu commands

Figure 5-4 on page 35 shows the Main Menu window. Each of the following commands brings you to a first level submenu:

- The `stats` menu gives statistics about the switch.
- The `cfg` menu contains all of the configuration options for the switch.
- The `oper` menu contains all of the operator commands. Some of these commands can change the state of the switch, but these changes only apply until the next reboot. They are not permanent.
- The `boot` menu contains the commands to control the booting of the switch, from which image to boot , which config to boot, and the **`gtimg`** and **`ptimg`** commands for getting and putting firmware files to the switch.

- The `maint` menu contains all of the commands for maintenance of the switch. The commands to manipulate the ARP cache and forwarding database are here, as well as the commands to obtain dumps of the current state of the switch for technical support.

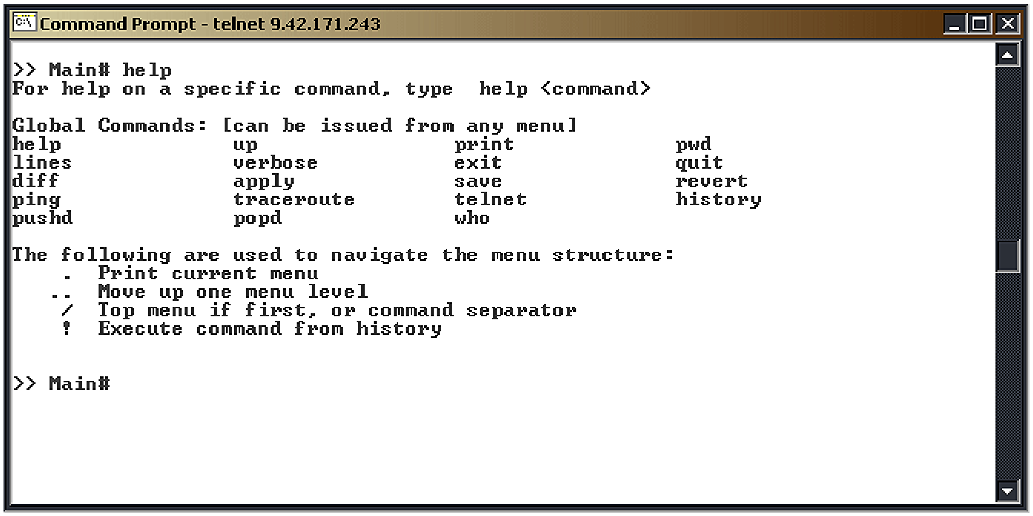


```
>> Main# /
------------------------------------------------------------
[Main Menu]
     info     - Information Menu
     stats    - Statistics Menu
     cfg      - Configuration Menu
     oper     - Operations Command Menu
     boot     - Boot Options Menu
     maint    - Maintenance Menu
     diff     - Show pending config changes  [global command]
     apply    - Apply pending config changes [global command]
     save     - Save updated config to FLASH [global command]
     revert   - Revert pending or applied changes [global command]
     exit     - Exit  [global command, always available]

>> Main# _
```

Figure 5-4 CLI Main Menu

Global commands

The remainder of the options on the Main Menu — `diff`, `apply`, `save`, `revert`, and `exit` — are all global commands that work anywhere on the switch. Figure 5-4 shows what each of the commands does. The `help` command is also global and lists all the global commands, as shown in Figure 5-5.



```
>> Main# help
For help on a specific command, type  help <command>

Global Commands: [can be issued from any menu]
help                 up                   print                pwd
lines                verbose              exit                 quit
diff                 apply                save                 revert
ping                 traceroute           telnet               history
pushd                popd                 who

The following are used to navigate the menu structure:
    .  Print current menu
   ..  Move up one menu level
    /  Top menu if first, or command separator
    !  Execute command from history


>> Main#
```

Figure 5-5 CLI global command list

Navigation commands

There are several commands that are useful in moving from one part of the menu tree to another. The commands are similar to those used in a UNIX® shell:

`cd` This command moves you to a given spot in the menu tree. Entering **`cd /`** always takes you back to the main menu.

`pwd` This command displays the current menu path where you are in the menu tree.

`up` This takes you back to the last menu that you touched.

`..` or **`cd ..`** Both of these commands take you up one level in the menu tree.

`pushd` and **`popd`** These commands allow you to manage a stack of menus that you visit frequently.

`history` This command displays the last several commands that you entered. You can reuse these commands by typing an exclamation point (**`!`**) followed by the number of the command as displayed.

`quit` or **`exit`** Both of these commands terminates your session.

Configuration control commands

These commands control the effectiveness of changes to the switch configuration. The general rubric for configuring the switch is EASY:

- E for editing the configuration, typing in your changes
- A for the **`apply`** command which makes the changes part of the running configuration
- S for the **`save`** command which writes the changed configuration to flash memory
- Y for yes, which is the answer to the prompt to be sure that you really want to update flash

Additional configuration control commands include the following:

`diff` This command displays the differences between the most recent edits and the running configuration.

`diff flash` This command displays the differences between the running configuration and its flash copy.

`revert` This command discards all changes which have not yet been applied.

`revert apply` This command discards applied changes which have not yet been saved to flash.

Additional commands

These are additional commands that facilitate troubleshooting or are otherwise helpful:

`ping` Sends ping, Internet Control Message Protocol (ICMP) echo, requests to the specified IP address.

`traceroute` Traces the IP path to a specified IP address.

`who` This command shows who is logged on to the switch and from which address.

`telnet` Opens a Telnet session to the designated IP address.

`verbose` Tailors the level of messages displayed on your session.

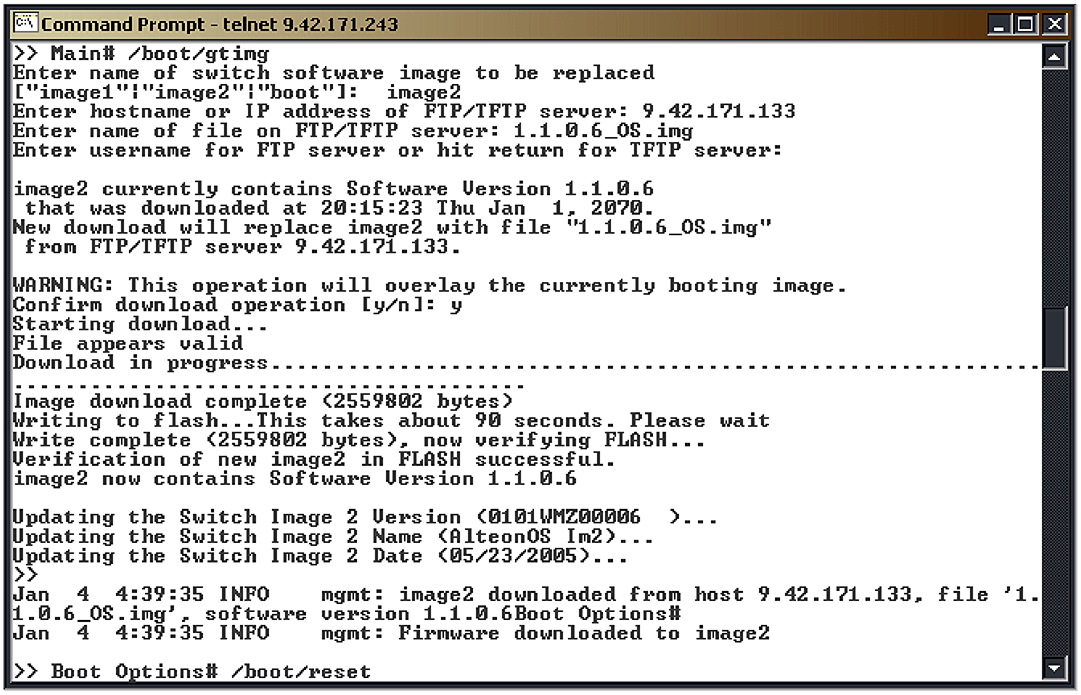
`lines` Controls the number of lines per screen for display purposes.

Upgrading the firmware

To upgrade the firmware on the Nortel Networks L2/3 GbESM, you must use Trivial File Transfer Protocol (TFTP) or File Transfer Protocol (FTP). It is not possible to use the Management Module menu item for upgrading firmware at this time. However, this is a planned feature for a future software release.

> **Important:** Before updating the firmware, save any configuration changes to the Nortel Networks L2/3 GbESM. From the Telnet session, enter `apply`**,** then press Enter**.** Type `save` and press Enter**.** Answer `y` to the prompt that asks to confirm saving to flash. Answer `y` to the prompt that asks if you want to change the boot to the active config block if it appears.

Figure 5-6 shows the process to load a new OS image file onto the switch.



```
>> Main# /boot/gtimg
Enter name of switch software image to be replaced
["image1"|"image2"|"boot"]:  image2
Enter hostname or IP address of FTP/TFTP server: 9.42.171.133
Enter name of file on FTP/TFTP server: 1.1.0.6_OS.img
Enter username for FTP server or hit return for TFTP server:

image2 currently contains Software Version 1.1.0.6
 that was downloaded at 20:15:23 Thu Jan  1, 2070.
New download will replace image2 with file "1.1.0.6_OS.img"
 from FTP/TFTP server 9.42.171.133.

WARNING: This operation will overlay the currently booting image.
Confirm download operation [y/n]: y
Starting download...
File appears valid
Download in progress............................................................
........................................
Image download complete (2559802 bytes)
Writing to flash...This takes about 90 seconds. Please wait
Write complete (2559802 bytes), now verifying FLASH...
Verification of new image2 in FLASH successful.
image2 now contains Software Version 1.1.0.6

Updating the Switch Image 2 Version (0101WMZ00006   )...
Updating the Switch Image 2 Name (AlteonOS Im2)...
Updating the Switch Image 2 Date (05/23/2005)...
>>
Jan  4  4:39:35 INFO    mgmt: image2 downloaded from host 9.42.171.133, file '1.
1.0.6_OS.img', software version 1.1.0.6Boot Options#
Jan  4  4:39:35 INFO    mgmt: Firmware downloaded to image2

>> Boot Options# /boot/reset
```

Figure 5-6 Display of a firmware update using CLI

The firmware for the Nortel Networks L2/3 GbESM is contained in two files: one is a boot image file and the other is the OS image file. Use the following steps to upgrade the firmware on the Nortel Networks L2/3 GbESM via the Telnet session:

1. Type **`/boot/gtimg`**.
2. Enter where the new image file will be placed. We are upgrading the boot image file, so enter `boot`. That is the location for the boot image file.
3. Enter the IP address of the TFTP server.
4. Enter the fully qualified path name for the boot image file that is on the TFTP server.
5. The switch reports the current version of the boot kernel on the switch and ask if you wish to replace it with new file. If you wish to continue, enter `y`.
6. When the download is finished, go back to Step 1, and repeat the process for the OS image file. In step 2, enter `image1` or `image2` as the location to store the new image file.

7. If the download location is the same as the location for the currently loaded OS image, the switch warns you that a failed download could result in an inoperative switch. If the download location is different from the location of the currently loaded OS image, the image file downloads. After the download is finished, the switch asks whether you want to use the old location or the new location. Figure 5-6 on page 37 shows a successful download of the OS image to image2.
8. Type **/boot/reset** to reset the switch and reboot with the new firmware files.

Capturing the current configuration

There are a few ways to capture the current configuration in the CLI. The first is to use a TFTP server to push the configuration file from the switch to the server. However, in some text editors the resulting file is a single long line of text. (We suggest using WordPad.) Although this method requires a TFTP server running in the network, it does work with any Telnet client. To capture the configuration by pushing a file to a TFTP server:

1. Enter **/cfg/ptcfg** at the command line.
2. Enter the IP address of the TFTP server.
3. Enter the filename to which you want to save the file.

A second way to capture the current configuration does not require a TFTP server. This method, however, requires a terminal emulator that can capture text. Example 5-1 uses a Windows® Telnet session to capture the text. The commands on the switch are the same for any software, but the steps to set the software to capture the text might be different. If your terminal emulator does not support this, you have to use the TFTP method. Using a Windows Telnet session and issuing the **/cfg/dump** command, the full switch configuration can be dumped.

Example 5-1 Example configuration file dump

```
>> Main# /cfg/dump
script start "Layer 2-3 Gigabit Ethernet Switch Module for IBM eServer BladeCenter" 4
/**** DO NOT EDIT THIS LINE!
/* Configuration dump taken 2:49:25 Sun Jan 4, 2070
/* Version 1.1.0.6, Base MAC address 00:11:f9:36:b7:00
/c/sys/access/user/uid 1
        name "USERID"
        pswd "727ac51410408000ba33a6f7d3f023f2186030e91e4bf6bc15dc8e028cfbe352"
        ena
        cos admin
/c/port INT1
        pvid 20
/c/port INT2
        pvid 10
/c/port INT3
        pvid 10
/c/port INT4
        pvid 99
/c/port EXT1
        tag ena
/c/port EXT2
        tag ena
/c/l2/vlan 10
        ena
        name "VLAN_Green"
        def INT2 INT3 EXT1 EXT2
/c/l2/vlan 20
        ena
        name "VLAN_Red"
        def INT1 INT2 EXT1 EXT2
/c/l2/vlan 99
        ena
        name "MGMT"
        def INT4 EXT1 EXT2
/c/l2/stg 1/clear
/c/l2/stg 1/add 1 10 20 99
/c/l2/lacp/port EXT1
        mode active
/c/l2/lacp/port EXT2
        mode active
        adminkey 17
/c/l3/if 99
        ena
        addr 10.99.0.243
        mask 255.255.255.0
        broad 10.99.0.255
        vlan 99
/c/l3/gw 1
        ena
        addr 10.99.0.245
/c/l3/gw 2
        ena
        addr 10.99.0.246
/
script end  /**** DO NOT EDIT THIS LINE!
```

Configuring user accounts

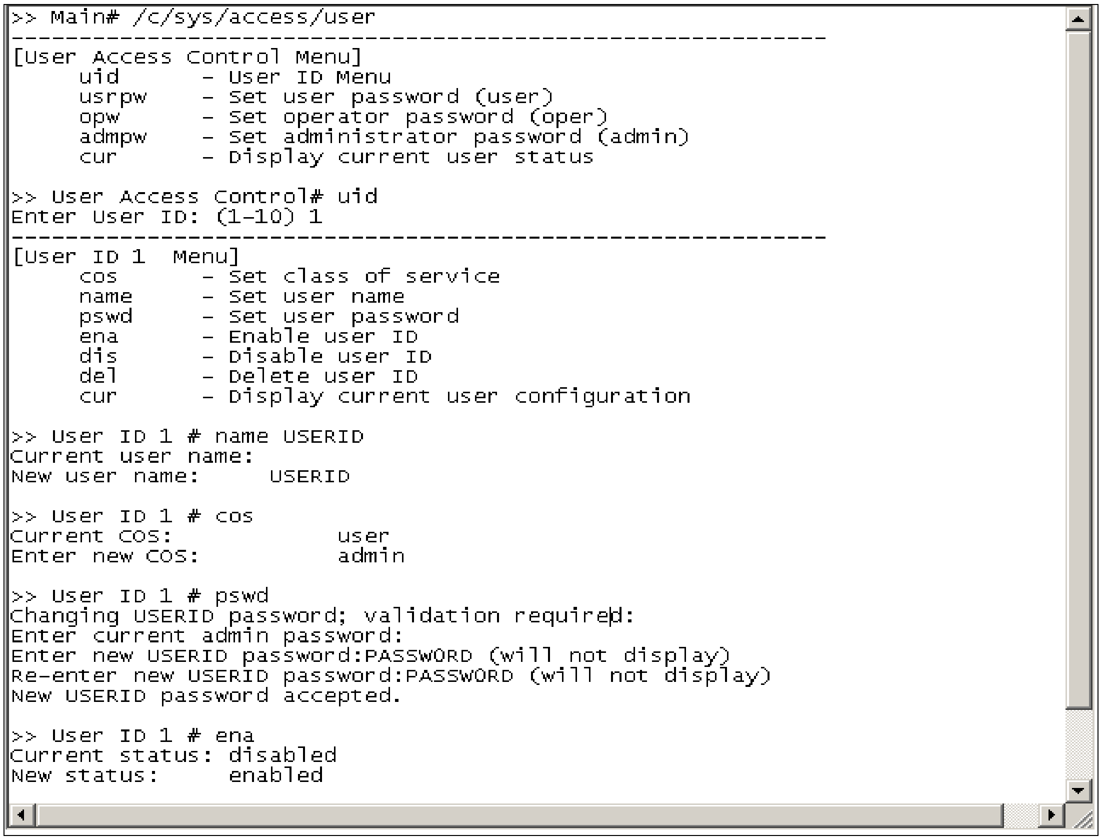
This section describes the user accounts on the switch. There are multiple modes of authentication which are supported on the Nortel Networks L2/3 GbESM:

- The default mode is to support passwords without individual user IDs. When accessing the CLI in this mode, there is only a single prompt to enter the password. Table 5-1 lists the three passwords that are supported.

Table 5-1 Description of default user accounts

User account	Description/Tasks performed	Default Password
User	Can view switch statistics but cannot make changes.	user
Operator	The Operator account manages all functions of the switch but cannot make permanent changes to the switch configuration.	oper
Administrator	Administrator is the super-user account and has complete access to all menus, information, and configuration commands on the switch.	admin

- Local mode allows the definition of individual user IDs with associated authority levels and passwords. This is configured in the `/cfg/sys/access/user` menu. For example, Figure 5-7 shows the configuration that is necessary to create the IBM @server BladeCenter default USERID and PASSW0RD account as an administrator.



```
>> Main# /c/sys/access/user
------------------------------------------------------------
[User Access Control Menu]
     uid       - User ID Menu
     usrpw     - Set user password (user)
     opw       - Set operator password (oper)
     admpw     - Set administrator password (admin)
     cur       - Display current user status

>> User Access Control# uid
Enter User ID: (1-10) 1
------------------------------------------------------------
[User ID 1  Menu]
     cos       - Set class of service
     name      - Set user name
     pswd      - Set user password
     ena       - Enable user ID
     dis       - Disable user ID
     del       - Delete user ID
     cur       - Display current user configuration

>> User ID 1 # name USERID
Current user name:
New user name:     USERID

>> User ID 1 # cos
Current COS:            user
Enter new COS:          admin

>> User ID 1 # pswd
Changing USERID password; validation required:
Enter current admin password:
Enter new USERID password:PASSWORD (will not display)
Re-enter new USERID password:PASSWORD (will not display)
New USERID password accepted.

>> User ID 1 # ena
Current status: disabled
New status:     enabled
```

Figure 5-7 Creation of USERID account as administrator of the switch module

- The third authentication mode supported on the Nortel Networks L2/3 GbESM is using an external authentication server. RADIUS and TACACS+ servers are both supported. These are configured in the `/cfg/sys/radius` and `/cfg/sys/tacacs+` menus respectively. More detail is available in the *Alteon OS 21.0 Application Guide*.

> **Note:** When you attempt to access the switch through the Web interface, you are prompted to enter a user name and password. For all the default users on the switch, the user name and password are the same string by default.

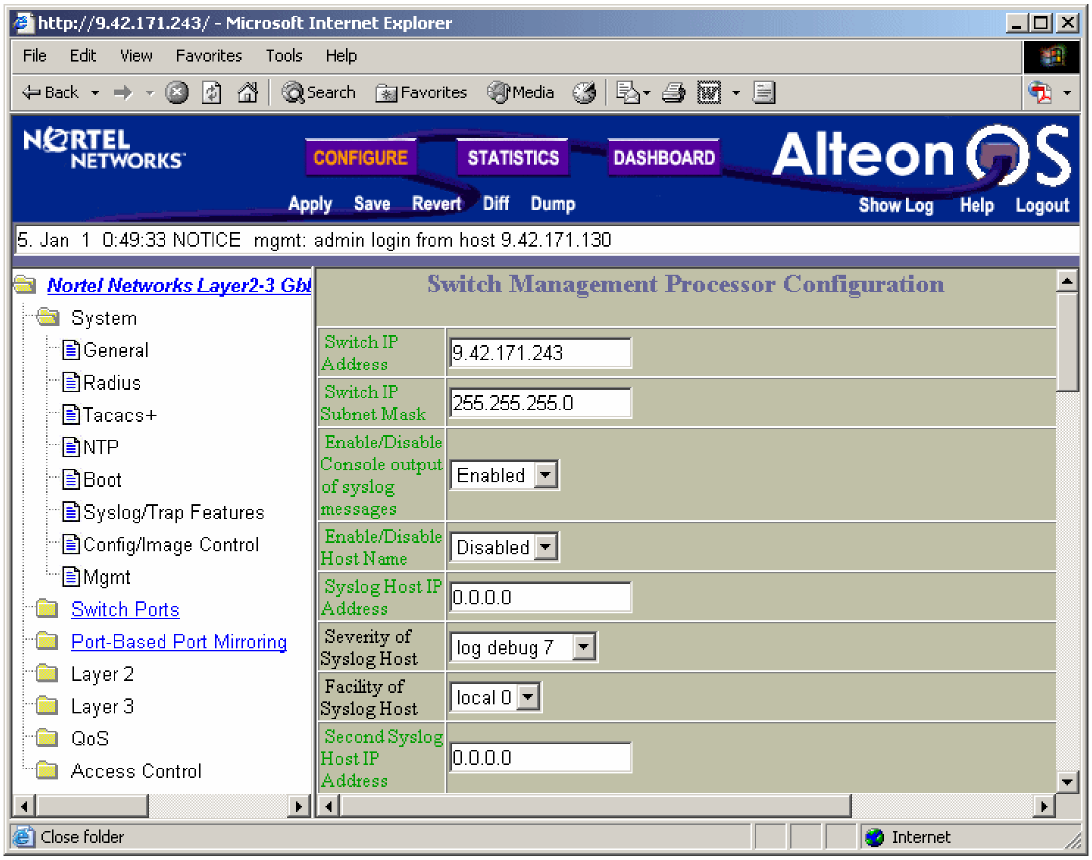
5.2.3 Browser Based Interface

We now take a brief look at the Browser-Based Interface (BBI) on the Nortel Networks L2/3 GbESM. Almost everything that can be done via the CLI can also be done in the BBI. In the remainder of this book, more emphasis is placed on configuring the switch using the CLI rather than using the BBI.

The Switch Information panel displays the MAC address of the switch as well as the firmware and hardware versions. Use the following steps to configure the system and contact information:

1. From the Nortel Networks L2/3 GbESM Web interface, click the folder icon next to Nortel Networks Layer 2/3 GbE Switch Module in the left-hand frame**.**

2. Click the folder icon next to System in the left-hand frame.
3. Click **CONFIGURE** at the top of the page.
4. Click the icon next to General in the drop-down list under System. On a window similar to Figure 5-8, you see options, such as IP Address and Network Mask fields, that can be configured on this page. Other options on this page include date and time settings, syslog settings (if you have a syslog server), and SNMP settings.


http://9.42.171.243/ - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
NORTEL NETWORKS
CONFIGURE STATISTICS DASHBOARD
Alteon OS
Apply Save Revert Diff Dump
Show Log Help Logout
Nortel Networks Layer2-3 Gbl
System
General
Radius
Tacacs+
NTP
Boot
Syslog/Trap Features
Config/Image Control
Mgmt
Switch Ports
Port-Based Port Mirroring
Layer 2
Layer 3
QoS
Access Control
Close folder
Internet

```
5. Jan  1  0:49:33 NOTICE  mgmt: admin login from host 9.42.171.130
```

Switch Management Processor Configuration

Switch IP Address	9.42.171.243
Switch IP Subnet Mask	255.255.255.0
Enable/Disable Console output of syslog messages	Enabled
Enable/Disable Host Name	Disabled
Syslog Host IP Address	0.0.0.0
Severity of Syslog Host	log debug 7
Facility of Syslog Host	local 0
Second Syslog Host IP Address	0.0.0.0

Figure 5-8 Switch information using BBI

You can browse through some of the other links in the left-hand frame to get more familiar with where the configuration options for the switch are located.

5. If you have made any changes to the switch and wish to save them, click **Apply** to apply the changes to the current running config.
6. Click **Save** to save the changes.

5.2.4 SNMP management - IBM Director

You can manage and monitor the Nortel Networks L2/3 GbESM switch module using SNMP via IBM Director. You can also use SNMP-based management systems, such as Tivoli Network Manager. The following SNMP capabilities are supported by the module:

- SNMP management stations can be configured to receive `TRAP` messages from the switch module. This is configured in the `/cfg/sys/ssnmp/` menu. Support is available for SNMPV3 as well as support for SNMP versions 1 and 2.
- SNMP Management Information Base (MIB*)* files are provided with every software image. These files can be imported to the MIB compiler, which is included with IBM Director and

other network management products. The MIBs that are provided include Nortel proprietary extensions to the standard MIB1 and MIB2 objects. Both read and write access to these variables can be configured.

5.3 Multiple Nortel Networks L2/3 GbESMs in a BladeCenter

If there are two (or more) switches in a single IBM @server BladeCenter chassis, the management (MGTx) interfaces of all of the switches are on VLAN 4095. This has the following consequences:

- All of the MGTx IP addresses that are configured through the Management Module Web interface should be on the same subnet as the Management Module internal and external port IP addresses (to allow for access through the Management Module). This configuration also makes it possible to Telnet from one switch module to another across the midplane of the chassis.
- It is not possible to pass substantive data between switch modules across the midplane using the MGTx ports. The Nortel Networks L2/3 GbESM will not forward data between the MGTx ports and any of the internal (INTx) or external (EXTx) ports. If you want to pass data from one switch module to another, then the modules must be either cabled directly to each other or connected by way of an external switch or router.

6

IBM @server BladeCenter system initial setup

This chapter discusses the network topology and the hardware that is configured to provide a tested and working configuration to help implement your Nortel Networks L2/3 GbESM for the IBM @server BladeCenter.

6.1 IBM @server BladeCenter system

In this section, we discuss the stages of preparing IBM @server BladeCenter for operation.

6.1.1 Management Module firmware

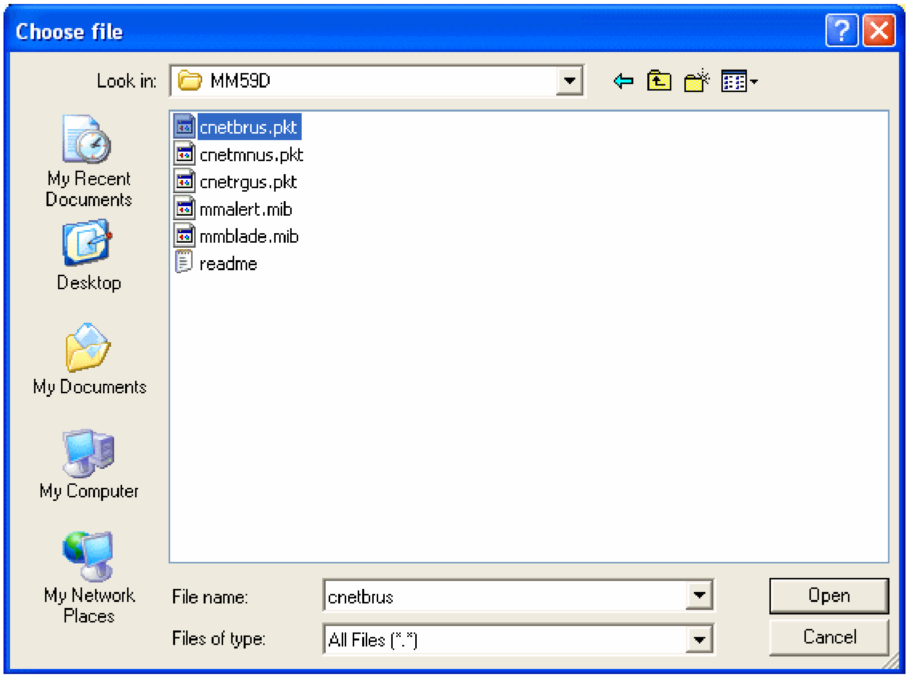
After the required hardware has been installed in your IBM @server BladeCenter, you should update the Management Module using IBM @server® BladeCenter - Management Module Firmware Update Version 1.18 or later. To acquire the firmware, go to:

`http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-54939`

Be sure to read the notes that are associated with the firmware installation. You can also search for the latest firmware version at:

`http://www-1.ibm.com/servers/eserver/support/bladecenter/index.html`

Follow the installation and setup instructions in the README file. You only need to install the files with the .pkt extension (see Figure 6-1). After the installation, you must restart the Management Module.



Look in: MM59D

cnetbrus.pkt
cnetmnus.pkt
cnetrgus.pkt
mmalert.mib
mmblade.mib
readme

File name: cnetbrus

Files of type: All Files (*.*)

Figure 6-1 Management Module firmware update files

6.1.2 Management Module network interface

In this section, we configure the Management Module external and internal network interfaces to exist upon the management subnet. The external network interface IP address is attached to the network outside of the IBM @server BladeCenter. This is the address used to contact the Management Module from an external device.

Establishing a physical connection to the Management Module

The only way to manage the Management Module is through the external 10/100 Mbps Ethernet port on the front of the module. To establish the physical connection to the Management Module, use one of the following methods:

- Use a Category 3, 4, 5, or higher unshielded twisted pair (UTP) straight-through cable to connect the Ethernet port on the Management Module to a switch in a network that has an accessible management station.
- Use a Category 3, 4, 5, or higher cross-over cable to connect a management station (PC or mobile computer) directly to the external Ethernet port of the Management Module.

Accessing the Management Module Web interface

After you establish the physical connection to the Management Module, configure the management station with an available IP address in the same subnet as the Management Module. By default, the subnet is 192.168.70.0/24. You have two primary methods to manage the Management Module:

- HTTP Web interface
- IBM Director

We use the Management Module Web interface to demonstrate the initial configuration of the Management Module and the switch module configuration.

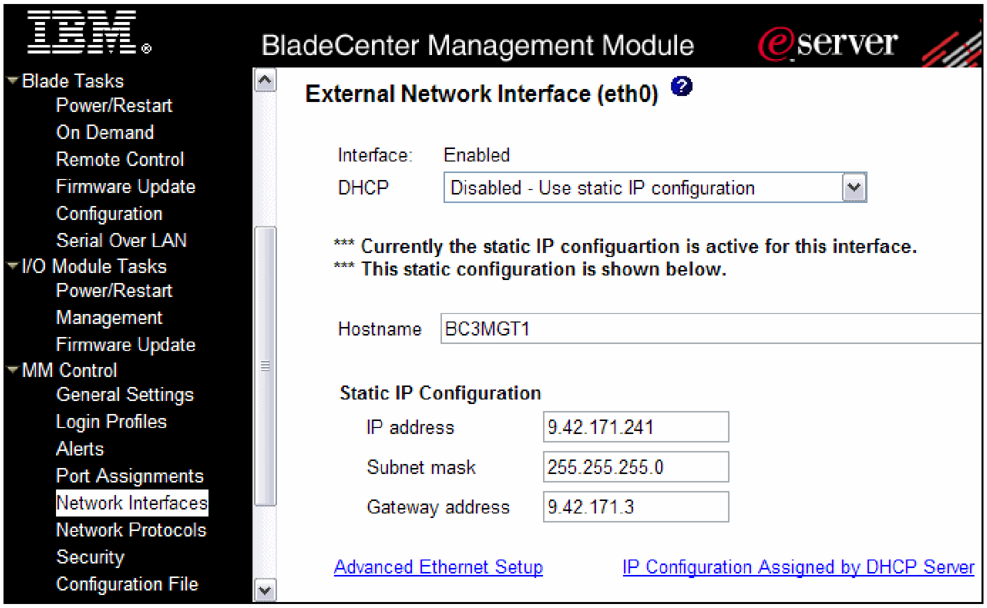
Follow these steps to establish a management session with the Management Module and to configure the initial switch module settings:

1. Open a Web browser and connect to the Management Module using the configured IP address. The default IP address for the Management Module external interface is 192.168.70.125. Note that the default IP address for the internal interface is 192.168.70.126.
2. Enter the user ID and password. The default is `USERID` and `PASSW0RD` (case-sensitive with a zero in the place of the letter O). Click **OK**.
3. At the initial window, click **Continue** to access the management session.

You can also refer to the *BladeCenter Management Module User’s Guide* on the IBM @server BladeCenter Documentation CD.

Configuring the Management Module network interfaces

After you access the Management Module Web interface, you can configure the external and internal network interfaces. From the BladeCenter Management Module Web interface, click **MM Control → Network Interfaces**.



External Network Interface (eth0)

Interface: Enabled

DHCP: Disabled - Use static IP configuration

*** Currently the static IP configuartion is active for this interface.
*** This static configuration is shown below.

Hostname: BC3MGT1

Static IP Configuration

IP address: 9.42.171.241

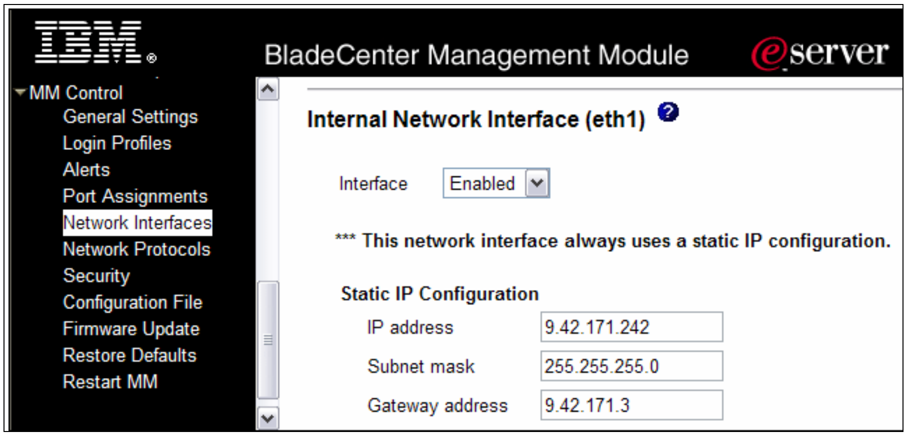
Subnet mask: 255.255.255.0

Gateway address: 9.42.171.3

Figure 6-2 Management Module External Network Interface window

The BladeCenter Management Module defaults to the IP address 192.168.70.125. If you have more than one BladeCenter chassis on your management network, you are required to change the external network interface (eth0). If you do not, you will have IP address conflicts and will not be able to access your Management Modules. In Figure 6-2, we configured the external interface to be on the same default management subnet with a unique IP address.

After you have configured the external interface, you need to configure the internal interface with another unique IP address. The purpose of internal network interface (eth1) is to communicate with the BladeCenter devices across an Ethernet link (Figure 6-3). Note that if you do not configure the internal interface on the same network as the external interface, you will not have IP connectivity from the Management Module to your switches modules.



Internal Network Interface (eth1)

Interface: Enabled

*** This network interface always uses a static IP configuration.

Static IP Configuration

IP address: 9.42.171.242

Subnet mask: 255.255.255.0

Gateway address: 9.42.171.3

Figure 6-3 Management Module Internal Network Interface window

Click **Save** at the bottom of the page. You must restart the Management Module to implement the changes.

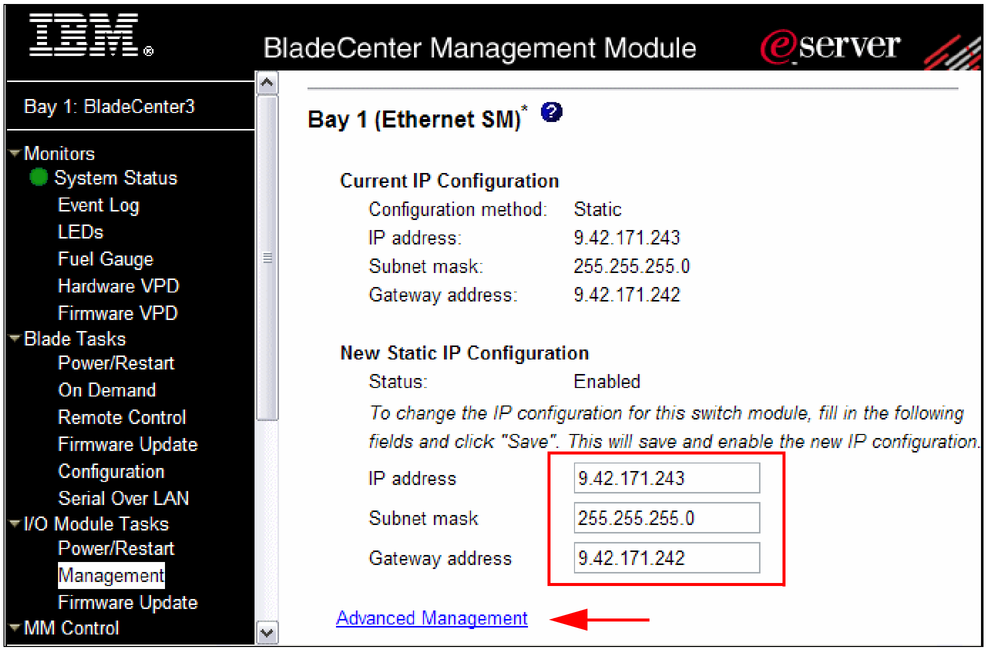
6.1.3 I/O module management tasks

In this section, we set up and configure the Nortel Networks Layer 2/3 Copper Gigabit Ethernet Switch Module for IBM @server BladeCenter.

Nortel Networks L2/3 GbESM setup and configuration

You can install the Nortel Networks L2/3 GbESM into any of the four BladeCenter switch bays in the rear of the chassis. Bay 1 is attached to one of the Ethernet Network Interfaces Controllers (NIC) on the blade HS20. Bay 2 is attached to the other Ethernet NIC. Each NIC is a Gigabit Full Duplex link to only one of the switches. As for HS40, which has a total of four NICs as standard, each two NICs link to one switch. A switch in bay 3 or bay 4 is required when a Gigabit Ethernet Expansion Card is being installed on the blade. This card provides an additional two NICs to the blades. One of the NICs has a dedicated Gigabit Full Duplex link to bay 3 and the other NIC to bay 4.

To manage the Nortel Networks L2/3 GbESM in bay 1, from the BladeCenter Management Module, click **I/O Module Tasks** → **Management**. A window similar to the one in Figure 6-4 opens.



Bay 1 (Ethernet SM)*

Current IP Configuration

Configuration method: Static
IP address: 9.42.171.243
Subnet mask: 255.255.255.0
Gateway address: 9.42.171.242

New Static IP Configuration

Status: Enabled

To change the IP configuration for this switch module, fill in the following fields and click "Save". This will save and enable the new IP configuration.

IP address 9.42.171.243
Subnet mask 255.255.255.0
Gateway address 9.42.171.242

Figure 6-4 I/O Module Tasks: Management (Bay 1 Ethernet SM) window

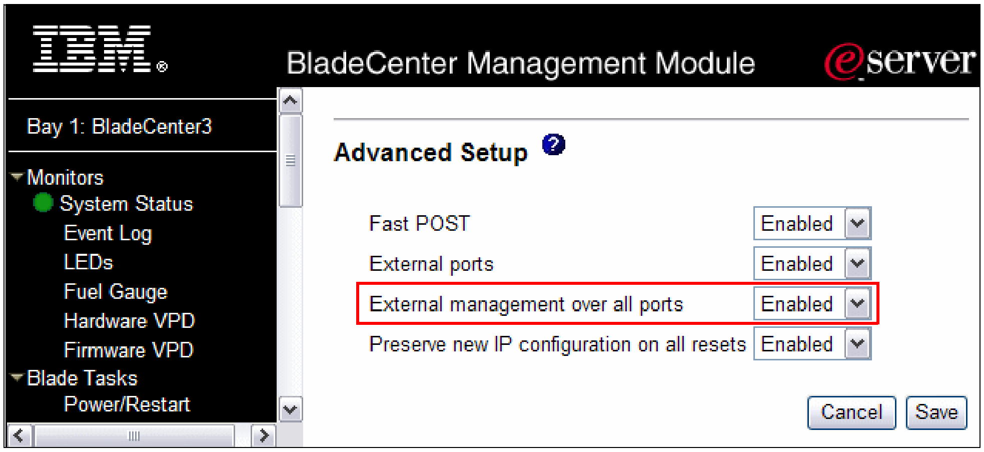
As with the Management Module, the switch must have a unique IP address and be on the same subnet as the Management Module for out-of-band management. Enter a Gateway address if attaching to other networks is required.

Click **Save** to apply these changes immediately. Rebooting or resetting is not required.

Enabling Nortel Networks L2/3 GbESM uplink ports through the Management Module

To enable the Ethernet ports of the Nortel Networks Layer 2/3 GbE Switch Module from the BladeCenter Management Module:

1. In the I/O Module Tasks → Management (Bay 1 Ethernet SM) window shown in Figure 6-4 on page 47, click **Advanced Management**.
2. If necessary, scroll down to the Advanced Setup section. You must at least set the External ports to Enabled for data to be sent out through the switch (Figure 6-5).
3. Click **Save** for the changes to be applied immediately.



Advanced Setup

Fast POST	Enabled
External ports	Enabled
External management over all ports	Enabled
Preserve new IP configuration on all resets	Enabled

Figure 6-5 I/O Module Tasks: Management - advanced setup

In our example, we enabled all options under Advanced Setup. Review the following list to determine which items you need to enable:

- Fast POST

  Use this field to enable or disable fast POST on this module. When fast POST is enabled, memory diagnostics are bypassed. When the field is disabled, memory diagnostics are executed during POST.

- External ports

  Use this field to enable or disable the external ports of this I/O module. When the external ports are disabled, no traffic can go through these ports. If this field is set to Disabled on the Management Module browser interface, the External ports will stay in disabled state as shown on the `/i/link` command no matter what commands are issued directly to the switch.

- External management over all ports

  Use this field to enable or disable external configuration management of this module. When this field is set to Disabled, only the Management Module ports can be used to change the configuration on this module (in other words, out-of-band management). When the field is set to Enabled, all ports (including internal, external, and Management Module ports) are enabled for management and you must follow certain rules.

- Preserve new IP configuration on all resets

  Use this field to specify whether you want the user-defined IP configuration to be preserved when the module's factory defaults are restored or when a reset is initiated by a source other than the Management Module. If this field is set to Enabled, be sure a valid IP configuration is entered for this switch module in the Management Module settings for this switch. If this field is set to Disabled, the factory default IP configuration will become active when the switch factory defaults are restored or when a switch reset is initiated by a source other than the Management Module. In this case, any user-defined IP configuration for the Nortel Networks L2/3 GbESM stored on the Management Module will not be used.

> **Note:** Although setting this value to Disabled allows the Nortel Networks L2/3 GbESM to use its stored IP information about subsequent reboots of the Nortel Networks L2/3 GbESM, when the Management Module reboots, it still places its version of the Nortel Networks L2/3 GbESM IP address on to the Nortel Networks L2/3 GbESM. Therefore, it is strongly recommended that you leave this setting at Enabled to prevent the different IP information from being used when the Nortel Networks L2/3 GbESM reloads, as opposed to when the Management Module reloads.

The only way to effectively use this setting as Disabled is to store the same information in the Management Modules Nortel Networks L2/3 GbESM settings as is stored on the Nortel Networks L2/3 GbESM. This ensures that no matter which reloads (the Management Module or the Nortel Networks L2/3 GbESM), the correct IP information is on the Nortel Networks L2/3 GbESM.

Nortel Networks L2/3 GbESM firmware download

In this section, we load the latest version of the switch module’s firmware.

Determining the level of Nortel switch software

After you install the Nortel Networks L2/3 GbESM in your BladeCenter unit, make sure that the latest Nortel switch operating system is installed on the module. To determine the level of the Nortel switch operating system software that is installed on the switch module:

1. Log on to the Nortel Networks L2/3 GbESM CLI.
2. Run **`/info/sys/general`** command.
3. Review the version information that is returned for the current revision.

Obtaining the latest level of switch software

To determine the latest level of the Nortel switch operating system software that is available from IBM, complete the following steps:

1. Go to the following Web address:

   `http://www.ibm.com/support/`
2. Click **Downloads and drivers**.
3. In the Downloads and drivers window Quick path field, enter the switch machine model number (for example, `8832-21x`) and click **Go**. A Results window opens, displaying a list of links to the latest available software.
4. Compare the level of software that you noted from the **`/info/sys/general`** command to the latest level of available software. If the two software levels do not match, download the latest level from the Web, and install it on your switch.

Upgrading the switch software

Switch software is upgraded through a TFTP server application. Typically, this software runs as an application under your operating system. Make sure that the software is installed on your server, then download the software images from the IBM Web site into a directory on your TFTP server. Enable the TFTP server and set its default directory to the one where the image is.

To transfer the software image files from the TFTP server to the switch, you must establish a Telnet session through the Management Module. To make sure that you have a connection, ping the TFTP server. The Telnet session performs optimally if all three network entities (TFTP server, Management Module, and switch IP addresses) are on the same subnet. Otherwise, you must use a router. Use the Management Module graphical interface to configure the IP addresses of the Management Module external network interface (eth0) and the Nortel Networks L2/3 GbESM so that they are on the same subnet as the TFTP server.

Note: Alternatively, you can use FTP instead of TFTP.

6.2 Blade server initial configuration

In this section, we prepare the IBM eServer™ BladeCenter HS20s for operation.

6.2.1 Firmware update

There are two primary methods to update the firmware of the BladeCenter HS20:

- Update diskettes

  Download the firmware diskette image. Create an update diskette, and boot the HS20 with it. The updates need to be done one at a time for each firmware.

- UpdateXpress CD

  IBM UpdateXpress provides an effective and simple way to update server firmware. UpdateXpress is a CD that contains a self-starting program which allows you to maintain your system firmware and Windows device drivers at the most current levels defined on the CD. UpdateXpress detects currently applied device driver and firmware levels automatically and presents them to you. It then gives you the option of selecting specific upgrades or allowing UpdateXpress to update all of the items that it detected as needing upgrades.

UpdateXpress

For our example, we used IBM UpdateXpress CD v4.01 - Servers to perform the firmware updates to our HS20 servers. To obtain UpdateXpress CD v4.01 - Servers, go to:

`http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-53046`

Note: Because updates are released as needed, UpdateXpress CD might not always provide the latest level updates at the time of configuration. This applies to both firmware and device drivers. Check the IBM Support Web site for updates that have been released later than the UpdateXpress CD:

`http://www.ibm.com/servers/eserver/support/xseries/index.html`

Getting started

Prior to starting IBM UpdateXpress, you should back up your system, unless you have a pristine system. The IBM UpdateXpress CD is a DOS-startable (bootable) CD. You can use the CD to start the system. You also can start the server from the hard disk drive and access files on the CD after the server starts.

Always update your system in the following order:

1. Update the device drivers. (Start from the hard disk and *access* the IBM UpdateXpress CD.)
2. Update the firmware. (Start from the UpdateXpress CD.)

Before the firmware update, make sure that your server can successfully restart.

> **Note:** In our example, we dealt with pristine HS20 systems. Therefore, we uploaded our firmware to the HS20 servers first. After loading the HS20s with the respective operating systems, we then launched UpdateXpress to update the operating systems with the supported device drivers.

Firmware update

In this section, we complete updating the firmware on the supported servers, HS20 Type 8832. To update the firmware:

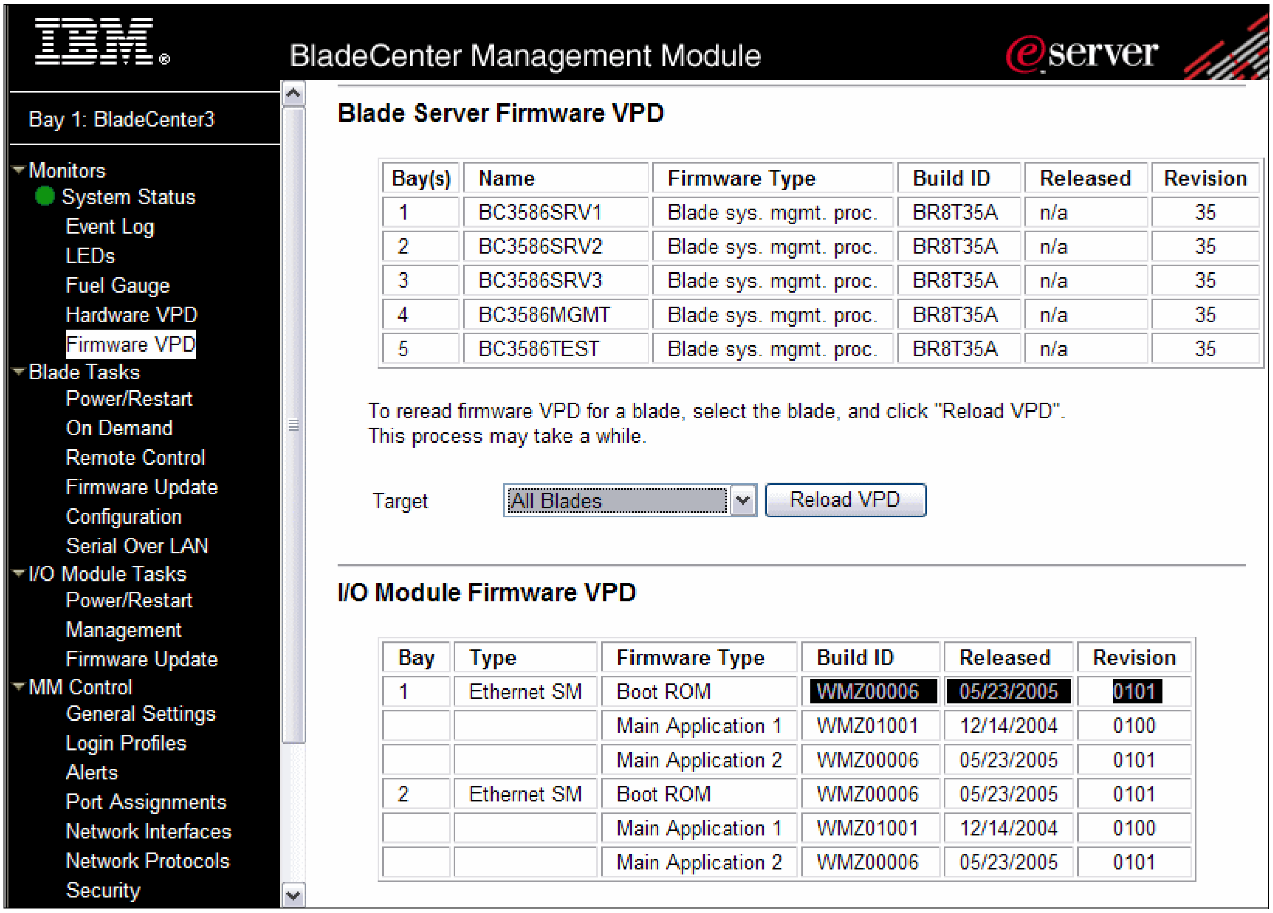
1. Start the system from the UpdateXpress CD.

   > **Note:** The Help button is not available in startable-CD mode. To view online help, go to the \help\Xpress directory on the UpdateXpress CD.
   >
   > All installed firmware components are displayed. If a firmware component needs to be updated or verified, it is selected automatically. If the firmware is at the same level as the firmware that is on the CD, the check box for that firmware is cleared.

   > **Note:** A 60-second countdown timer is displayed in the Firmware Update window. The selected firmware components are updated automatically when the timer reaches zero. To stop the timer, press any key.

2. Select or deselect the firmware components to be updated.
3. Click **Apply Update**.
4. Remove the UpdateXpress CD from the CD-ROM drive. Then, restart the server.

After UpdateXpress completes updating the firmware to your servers, and so forth, you should review current firmware levels by selecting **Monitors** → **Firmware VPD** on the Management Module Web interface, which opens a window similar to the one shown in Figure 6-6 on page 52.



Blade Server Firmware VPD

Bay(s)	Name	Firmware Type	Build ID	Released	Revision
1	BC3586SRV1	Blade sys. mgmt. proc.	BR8T35A	n/a	35
2	BC3586SRV2	Blade sys. mgmt. proc.	BR8T35A	n/a	35
3	BC3586SRV3	Blade sys. mgmt. proc.	BR8T35A	n/a	35
4	BC3586MGMT	Blade sys. mgmt. proc.	BR8T35A	n/a	35
5	BC3586TEST	Blade sys. mgmt. proc.	BR8T35A	n/a	35

To reread firmware VPD for a blade, select the blade, and click "Reload VPD".
This process may take a while.

Target: All Blades

I/O Module Firmware VPD

Bay	Type	Firmware Type	Build ID	Released	Revision
1	Ethernet SM	Boot ROM	WMZ00006	05/23/2005	0101
		Main Application 1	WMZ01001	12/14/2004	0100
		Main Application 2	WMZ00006	05/23/2005	0101
2	Ethernet SM	Boot ROM	WMZ00006	05/23/2005	0101
		Main Application 1	WMZ01001	12/14/2004	0100
		Main Application 2	WMZ00006	05/23/2005	0101

Figure 6-6 BladeCenter Firmware VPD window

6.2.2 Operating systems

In this section, we prepare the use of our operating systems for the BladeCenter HS20s. Note you can use IBM Director and Remote Deployment Manager (RDM) to customize and deploy your network operating systems to the HS20s. However, if you are building your network operating system manually, after Windows 2003 Server has been successfully installed, go the next section 6.2.3, “Broadcom Advanced Control Suite installation” on page 53.

Microsoft Windows 2003 Broadcom driver installation

Windows 2003 does not ship with the drivers needed to run the Broadcom Ethernet NICs. The drivers must be updated for the NICs to be usable.

To obtain the Broadcom NetXtreme Gigabit Ethernet drivers for your Microsoft Windows 2003 systems, go to:

http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-43815

For our example, this Web site provided us with the Broadcom NetXtreme Gigabit Ethernet Software CD for the BCM570x-based servers and adapters Version 8.1.6. You should acquire Version 8.1.6 or later for setting up your operating system environment. This supports the following machines:

- IBM @server BladeCenter HS20 (Type 8678, 8832, 8843)
- IBM @server BladeCenter JS20 (Type 8842)
- IBM @server BladeCenter LS20 (Type 8850)

- IBM @server 325, 326
- IBM @server MXE-460
- IBM @server xSeries 205, 225 (Type 8647), 226, 235, 236, 255, 305, 335, 336, 346, 365, 440, 445, 450, 455, 460
- IBM IntelliStation A Pro (Type 6224)
- IBM IntelliStation E Pro (Type 6216, 6226)
- IBM IntelliStation Z Pro (Type 6221, 6223, 6227)
- IBM IntelliStation M Pro (Type 6219, 6225, 6228)

Red Hat Linux Broadcom driver installation

To perform a driver installation for Red Hat Linux®, use the example Red Hat Linux AS 2.1 Broadcom driver installation in Chapter 6 of the Cisco Systems Intelligent Gigabit Ethernet Switch Module for IBM @server BladeCenter that is available at:

http://www.redbooks.ibm.com/redpapers/pdfs/redp3869.pdf

6.2.3 Broadcom Advanced Control Suite installation

Network interface card (NIC) teaming is one method for providing high availability and fault tolerance in IBM @server servers. In this example, we use Broadcom Advanced Services Program (BASP) to implement teaming functionality along with load balancing, fault tolerance, and VLAN tagging.

To enable NIC teaming, the Broadcom Advanced Control Suite (BACS) application must be used on the HS20s. The program is included with the drivers, which you can download at:

http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-43815

To install the suite:

1. Navigate to the location where the Broadcom Advanced Control Suite application files were extracted (default C:\Drivers\Bcom*XXX*, where *XXX* is the code level). Execute Launch.exe. You see a window similar to the one shown in Figure 6-7 on page 54.

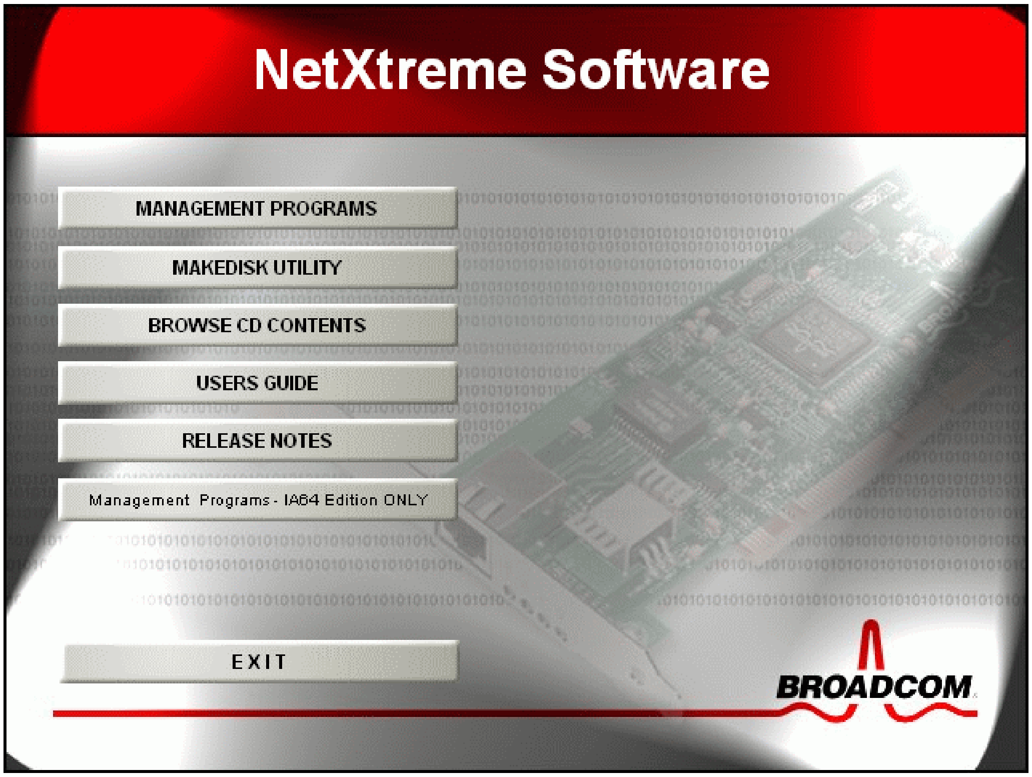
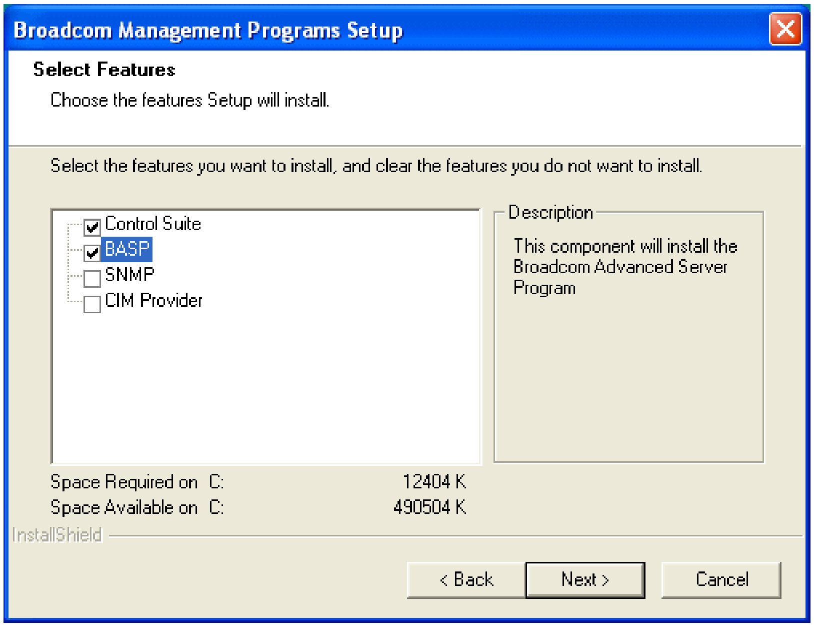


Figure 6-7 Broadcom selection window

2. Click **MANAGEMENT PROGRAMS**, and a window similar to Figure 6-8 opens.



Select Features

Choose the features Setup will install.

Select the features you want to install, and clear the features you do not want to install.

☑ Control Suite
☑ BASP
□ SNMP
□ CIM Provider

Description

This component will install the Broadcom Advanced Server Program

Space Required on C: 12404 K
Space Available on C: 490504 K

Figure 6-8 Select Features window

3. Select **Control Suite** and **BASP**.
4. Click **Next** to continue, and then click **Finish**.

6.3 Firmware and device drivers used in this example

We applied the following firmware and drivers to our environment:

- IBM @server BladeCenter Management Module:
  - Management Module Firmware Update Version 1.10
- BladeCenter HS20(8832) firmware:
  - BladeCenter HS20 (Type 8832) - Flash BIOS Update Version 1.09
  - BladeCenter HS20 (Type 8678, 8832) - blade server integrated system management processor firmware update Version 1.09
  - Broadcom NetXtreme firmware level 3.21
- Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter firmware:
  - Nortel Networks L2/3 GbESM firmware build level 1.0.1.6
- BladeCenter HS20(8832) device drivers for Windows 2003 Advanced Servers:
  - Broadcom NetXtreme Device Driver 8.22.1.0
  - Broadcom Advanced Server Program 8.1.4
  - Broadcom Advanced Control Suite 8.1.4

Visit the following Web site for the latest software and drivers:

http://www-307.ibm.com/pc/support/site.wss/DRVR-MATRIX.html

7

Nortel Networks L2/3 GbESM configuration and network integration

This chapter discusses the Best Practices for implementing and configuring Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules in Nortel Networks, Cisco Systems, and Extreme Networks network environments. It provides several network topology examples to help you successfully implement the Nortel Networks Layer 2/3 GbE Switch Module.

7.1 Standards and technologies

This section provides a brief overview of the networking standards which are supported on the Nortel GbESM. Detailed examples of configurations which include these standards are included later in this chapter.

The bulk of this chapter shows interconnection with Cisco Systems' devices. In 7.9, "Configuration for Extreme switches" on page 143, we include configurations for some functions that were tested on a pair of Extreme switches. Interconnection of the GbESMs with Nortel switches which run Alteon OS software is easily accomplished because the command syntax is almost exactly the same as that of the GbESMs themselves.

The Nortel GbESM connects to and interoperates with products from a variety of vendors which fully implement these standards. A limited number of other vendors' proprietary extensions to the standards are also supported. The Nortel GbESM configurations provided in this chapter are not limited to use when connecting with Nortel, Cisco, or Extreme gear.

7.1.1 VLAN tagging - 802.1Q

This standard defines the use of a tag field in the header of each packet which identifies the VLAN to which the packet belongs. This feature is configured with the **`/cfg/port/tag ena|dis`** command.

Cisco refers to this as a *trunk* and uses the **`switchport trunk`** command to configure it. On the Nortel Networks L2/3 GbESM, the internal ports have tagging enabled by default to provide support for Serial over LAN functionality. The external ports have tagging disabled.

7.1.2 Link Aggregation and LACP - 802.3ad and 802.3-2002

These standards define techniques for grouping two or more parallel connections between a pair of devices as a single logical link or trunk with the total bandwidth of its members. The Nortel GbESM supports a prestandard implementation of static trunking which is interoperable with similar implementations from Cisco and Extreme, among others. This is configured with commands on the **`/cfg/l2/trunk`** menu.

The Nortel GbESM also supports Link Aggregation Control Protocol (LACP) as defined in the standards. LACP provides dynamic negotiation of the formation of trunks and ensures that the two devices' configurations are consistent. This feature is configured from the **`/cfg/l2/lacp`** menu.

Note that Cisco refers to this as a *Port Channel* or *EtherChannel* and uses the **`channel-group`** and **`interface portchannel`** commands to configure it. Cisco also supports a proprietary prestandard protocol known as Port Aggregation Protocol (PAgP) as well as supporting LACP. PAgP and LACP provide almost identical functionality.

7.1.3 Spanning Tree - 802.1D, 802.1w, 802.1s

These protocols define techniques for managing Layer 2 networks whose topologies include loops. They ensure that a broadcast packet (or any other packet) is not forwarded endlessly around such a loop by logically blocking some ports.

Original or Classic Spanning Tree Protocol (STP) is defined by the 802.1D standard. This standard is relatively old and many vendors, including Cisco, Nortel, and Extreme, have implemented proprietary extensions to it. The Nortel GbESM will interoperate with Cisco's proprietary Per VLAN Spanning Tree (PVST). An example of this is shown in 7.7, “Advanced Layer 2 topology sample configurations” on page 80.

Two of the key shortcomings of the original STP standard is that it takes as much as 50 seconds to recover from the failure of a link or device, and that it does not deal well with multiple VLANs carried over the same physical link (typically by using the 802.1q standard described above.) These shortcomings are remedied by the 802.1w standard for Rapid Spanning Tree protocol (RST or RSTP) and by the 802.1s standard for Multiple Spanning Tree or Multiple Instance Spanning Tree protocol (MSTP or MISTP).

The Spanning Tree functions are configured on the Nortel GbESM with commands from the **`/cfg/l2/stg`** and **`/cfg/l2/mrst`** menus.

> **Restriction:** In our testing, the Cisco implementation of MSTP would only work with other Cisco products and produced undesirable results when connected to other vendors’ products. The Ethereal packet decoder was unable to completely decode the BPDUs sent by the Cisco switches. After our testing was completed, we learned that this is remedied in IOS version 12.2(25)SEC on the Cisco 3750. We assume that similarly numbered versions on other platforms will also include this fix.

7.1.4 Routing Information Protocol - RFC1058 and RFC2453

RIP is used by Layer 3 routers to exchange routing table information about the networks which they can reach and determine how far away are those networks. This facilitates end-to-end IP connections which traverse multiple routers. Servers can be configured to listen to RIP information but this is rarely done.

RFC1058 defines the original RIP specification, which was enhanced with the definition of RIP version 2 (RIP2) in RFC 2453. RIP version 1 has significant shortcomings compared to RIP2, and therefore RIP1 has largely fallen out of favor. These shortcomings include limitations on the size of network which can be supported and the completeness of the route information which can be exchanged using RIP1.

All of the products tested support RIP V2. To configure RIP on the Nortel GbESM, use commands on the **`/cfg/l3/rip`** menu.

7.1.5 Open Shortest Path First (OSPF) - RFC1257, RFC2328, and others

OSPF, like RIP, is used by Layer 3 routers to exchange routing table information. It is more scalable and versatile than RIP and recovers from failures more quickly. However, OSPF is also more complex and more difficult to configure.

OSPF uses a fundamentally different approach to managing routing tables than RIP in that each router running OSPF maintains a complete representation of the network topology; with RIP, routers are only aware of their immediate neighbors. As a result of this, OSPF requires more memory and more processing power than RIP.

All of the products tested support version 2 of OSPF. To configure OSPF on the GbESM, use commands on the **`/cfg/l3/ospf`** menu.

7.1.6 Virtual Router Redundancy Protocol (VRRP) - RFC 3768

Virtual Router Redundancy Protocol is used by Layer 3 routers to enable one (or more) routers to back up a primary or master router seamlessly. Workstations and other devices are typically not aware when a backup router takes over for a primary router which has failed. VRRP recovery time can be as little as one second, or less.

VRRP is used by the Nortel GbESM to facilitate the implementation of High Availability designs using the IBM @server BladeCenter. Sample configurations which use VRRP are included in 7.8, “Layer 3 topology sample configurations” on page 108. The Cisco and Extreme switches, and other devices which are upstream from the Nortel GbESMs do not have to support VRRP in order to interoperate with the Nortel modules. (However, Extreme does support VRRP and Cisco supports a proprietary protocol called HSRP which is very similar.) There is some driver configuration which must be done on the server blades to set the default gateway of the server to point to the VRRP address.

To configure VRRP on the Nortel GbESM, use the **`/cfg/l3/vrrp`** menu.

7.1.7 Where standards originate and how to get them

The below is for informational purposes only. It is not necessary to read the standards to successfully configure the technologies they define.

Standards identified as part of the 802 series are from the Institute of Electrical and Electronic Engineering (IEEE). You can find them at:

`http://standards.ieee.org/getieee802`

Some newly approved or draft standards require the payment of a fee to obtain the text of the standard.

Standards identified as RFCs come from the Internet Engineering Task Force (IETF). There are multiple Web sites where all or some of the RFCs can be downloaded without charge. One such site is the following:

`http://www.ietf.org`

7.2 Summary of sample configurations

The remaining sections in this chapter present several sample configurations which exploit different capabilities of the Nortel Networks L2/3 GbESM. It is not intended that any of these samples be copied exactly and used in a real network. However, portions of one or more of them can serve as the basis for a real configuration.

7.2.1 Basic Layer 2 configuration

This configuration is the basis for all of the other samples. It includes connections between two Nortel GbESM modules and two upstream Core switches using link aggregation and VLAN tagging. This configuration provides sufficient redundancy to protect against some single point failures but not as many as those which follow.

The configurations tested include the use of *trunk failover*, a feature of the Nortel GbESM software which allows an application to survive failures of the links between the GbESM and upstream switch(es) or failures of the upstream switch(es) themselves. Use of this feature enables the design of a robust High Availability configuration. Note that it is possible to use trunk failover in topologies like those of the advanced Layer 2 configurations in the following section. You can find more information about this configuration in 7.6, “Basic Layer 2 entry topology” on page 69.

7.2.2 Advanced Layer 2 configurations

These configurations add additional connections between the GbESM modules and the Core switches, using a mesh topology. This topology includes loops on most of the VLANs used, and therefore requires the use of Spanning Tree Protocol (STP). Testing was performed using the original STP (802.1D) as well as Rapid Spanning Tree and Multiple Spanning Tree (802.1w and s). You can find more information about this configuration in 7.7, “Advanced Layer 2 topology sample configurations” on page 80.

7.2.3 Layer 3 configuration - static routing

This configuration uses the Layer 3 IP routing capabilities of the Nortel GbESM. Not only are the blade servers on differing VLANs from each other, in this configuration the connections to the Core switches are on different VLANs from all of the blades. Thus, every packet sent to or from the blades is routed at Layer 3.

This configuration includes a full mesh topology, connecting the each of the two GbESM switches and the two Core switches to all of the others. STP is not needed in this design despite the full mesh because of the use of Layer 3 routing.

High Availability is achieved in this design through the use of Virtual Router Redundancy Protocol (VRRP) and the Hot Standby option. VRRP allows the two GbESM modules to back each other up, so that the failure of one GbESM can be survived. Hot Standby enhances the capabilities of VRRP by allowing the backup switch to take over in the event the primary switch is cut off from its upstream neighbors.

One key issue with Layer 3 designs is the need for L3 routers to know where to forward traffic to a given destination. In this design, both the Core switches and the GbESM switches are explicitly configured with the IP addresses of their neighbors. This approach is referred to as *static routing*, and does not itself react to changes in the network topology. Topology changes such as IP address reassignments, introduction of additional devices or networks, and others could require configuration changes on all of the switches in our test environment.

You can find more information about this configuration in 7.8, “Layer 3 topology sample configurations” on page 108.

7.2.4 Layer 3 configurations - dynamic routing

These configurations are similar to the ones in the previous section but they use *dynamic routing* protocols instead of static routing. These protocols enable the GbESM modules and Core switches to provide each other with up to date information about which IP subnets they can reach, allowing traffic to be forwarded on the optimal path. In addition, these dynamic protocols react quickly to changes in network topology and do not require that every switch have its configuration updated when the network changes.

These configurations include the use of VRRP — without hot-standby — on the portion of the network connecting directly to the server blades. Hot standby can be used in a configuration such as this if desired.

The decision to use static or dynamic routing is made by network architects in most cases. These configurations are provided to show that the GbESM switch modules can participate in RIP or OSPF networks if it is decided that this is the best way to integrate them with the existing network. You can find more information about this configuration at 7.8, “Layer 3 topology sample configurations” on page 108.

7.3 Introduction to High Availability

This section provides an explanation of the trunk failover feature, the Broadcom Advanced Services Protocol driver, and VRRP and of how they work together to provide a highly available IBM @server BladeCenter environment.

7.3.1 Introduction to trunk failover

Trunk failover works by shutting down ports directly connected to the configured blade servers when the configured upstream trunks go down. The internal ports are put into *disabled* state, and the servers react as though the cable to the network card on a free-standing server had been unplugged. When the configured external trunks recover, the internal ports are re-enabled.

Trunk failover is intended to prevent the following failure mode, when used as part of a High Availability design (Figure 7-1 on page 63):

- The critical connections between a Nortel GbESM and upstream switch(es) fail, due to a cable problem or the failure of the upstream switch(es).
- The Nortel GbESM continues to function, and the server blades continue to send traffic to it.
- The Nortel GbESM, having lost its upstream connections, has no place to forward the server blades’ traffic and therefore discards it.

The Nortel GbESM also supports a feature called *Hot Standby* which provides similar function to trunk failover. However, Hot Standby can only be used in a Layer 3 configuration in concert with VRRP.

Note that if the Nortel GbESM itself fails, High Availability can be provided through the use of other features such as NIC teaming and VRRP.

Configuration

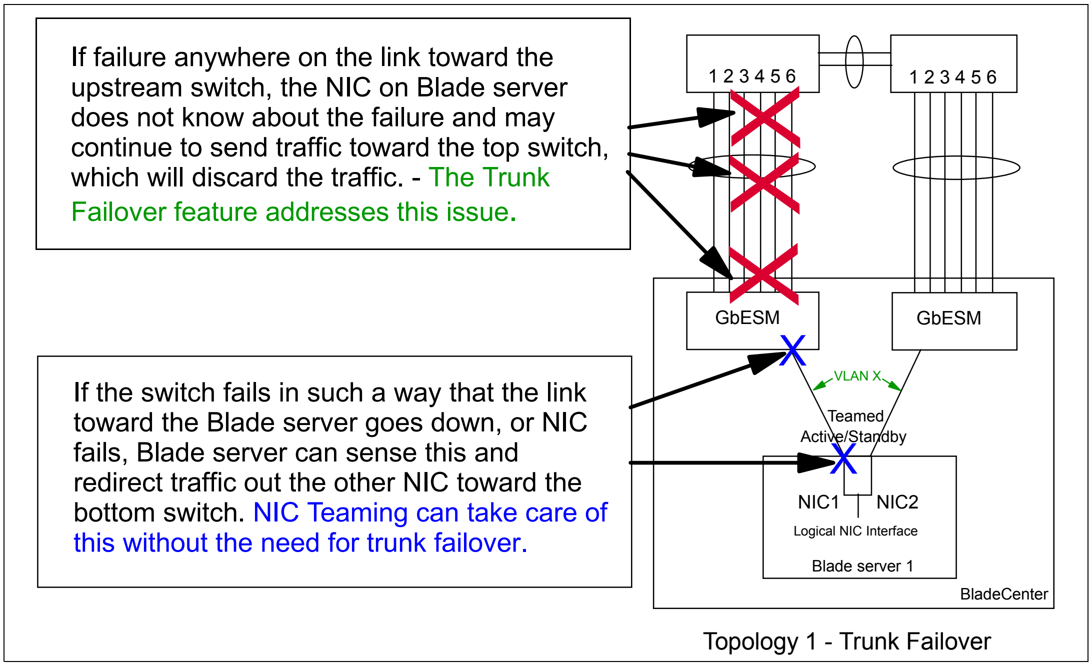
Trunk failover is configured on the Nortel GbESM with the **`failover ena|dis`** command, as follows:

```
/cfg/l2/trunk 1
failover ena
```

If there are multiple trunk groups which are critical upstream connections — such as to multiple upstream switches — then they should all have the failover feature enabled. Failover will not occur until all of them fail at the same time.

In most cases, you should configure trunk failover on all Nortel Networks L2/3 GbESM in the IBM @server BladeCenter if the server blades are running NIC Teaming. These two features work together to provide a High Availability design.

Restriction: The currently available release (1.0.1.6) of software for the Nortel Networks Layer 2/3 Copper and Fiber GbE Switch Modules for IBM @server BladeCenter does not support trunk failover for trunks configured with LACP. This feature is to be added in a forthcoming release. This results in a slight change in the command syntax required. We were able to validate this briefly with an early test version of the next release of software.



If failure anywhere on the link toward the upstream switch, the NIC on Blade server does not know about the failure and may continue to send traffic toward the top switch, which will discard the traffic. - The Trunk Failover feature addresses this issue.

If the switch fails in such a way that the link toward the Blade server goes down, or NIC fails, Blade server can sense this and redirect traffic out the other NIC toward the bottom switch. NIC Teaming can take care of this without the need for trunk failover.

Figure 7-1 What trunk failover can protect against

7.3.2 Introduction to NIC Teaming

NIC Teaming is a function that is provided by Broadcom, the manufacturer of the NIC chips used on the Blade Servers, in their software. Broadcom provides the Broadcom Advanced Services Protocol (BASP) which includes NIC teaming, as well as the Broadcom Advanced Control Suite (BACS) which is a Windows application which helps configure NIC teaming.

NIC teaming allows two or more physical NICs to be treated as a single logical network object in Windows or a single /dev/eth file in Linux. The single object or file can then be assigned network properties such as an IP address in the same way as any other NIC.

The BACS application allows several types of teams to be created. For HA designs, the Smart Load Balancing (SLB) team is used. Layer 2 designs can have both of the adapters (on an HS20 blade) as active members of the team; for Layer 3 designs, an active or standby team is used with one adapter as an active member of the team and the second adapter as a standby member of the team.

NIC teaming is intended to provide both additional capacity (bandwidth) as well as High Availability. The team will detect loss of signal on any of its member NICs and continue to send traffic through other active members, or activate standby members if necessary. In the IBM @server BladeCenter, NIC teaming will detect the failure of a NIC chip on the server blade, the loss of connection to a switch module via the midplane, and the failure of a switch module (including intentional removal or power-off). Of these, intentional removal or power-off of a switch module is by far the most common.

The BASP drivers also provide support for 802.1q tagging of the server NIC. This allows support for multiple VLANs on a single physical NIC or on a group of teamed NICs. When this capability is used, each VLAN has its own network object (windows) or /dev/eth file (Linux). Thus, each VLAN can be assigned its own IP address. This can be useful to isolate different categories of traffic from each other or to provide different Quality of Service (QoS) configurations for different types of traffic whose target is the same server. A sample configuration of this capability is included in 7.5.3, “Base configuration common to all examples” on page 68.

Notes:

- The BASP driver can be configured to use standards-based Port Aggregation (802.3-ad) teaming. This is useful on HS40 blades or HS20 blades with the SCSI sidecar, both of which have two ports connecting them to each switch module. Only ports connected to the same switch should be teamed in this way.
- The current production version of the GbESM software(1.0.1.6) does not support trunking on internal ports. The next (1.1) software release will add this function.
- The Nortel L2/7 GbESM (but not the L2/3 GbESM) supports a capability called Server Load Balancing (SLB). This is not similar to the Broadcom Smart Load Balancing; it involved multiple servers running the same application.
- Some of the previous descriptions contain *as is* information based on a test in our specific environment with BASP 7.12.01, the latest as this paper is written, and might differ in different environments or future software releases.

For more information about BASP NIC teaming, refer to the BACS online help and *BCM570X Broadcom NetXtreme Gigabit Ethernet Teaming* white paper, which is available at:

http://www.broadcom.com/collateral/wp/570X-WP100-R.pdf

7.3.3 Introduction to VRRP

Virtual Router Redundancy Protocol (VRRP) is a Layer 3 protocol used to enable switches to back each other up in a way which is transparent to client and server computers. VRRP works by defining an address which is shared between the switches. One switch which is the **Master** is the only one which will answer to the shared address. One or more other switches in **Backup** state are configured to take over from the master in the event of a failure. An instance of VRRP is configured for each VLAN where a shared address is to be used. This implies that if there is one VLAN for the internal ports and an additional VLAN for the external ports, then there can be two instances of VRRP, providing a shared address on the internal VLAN and a different shared address on the external VLAN. An example configuration which illustrates this is in 7.7, “Advanced Layer 2 topology sample configurations” on page 80.

VRRP Priority

Each switch in a group running VRRP has a configured priority. When VRRP first becomes active, the switch with the highest priority will become the Master switch. The master switch sends out periodic **hello** packets announcing that it is still operational. The backup switch with the highest configured priority will take over when the hello packets are no longer received.

There are configuration options, called *tracking* options, which adjust the priority of a switch dynamically based on the number of certain categories of resources (such as ports) which are available. Use of these options can allow a backup switch to take over even if the current master is still running but has lost some of the tracked resources.

7.3.4 Some important rules for ensuring High Availability

For High Availability (HA) to be truly effective it needs to be well thought out. A complete High Availability design should encompass servers, storage, and more of the network than just the portions connected to the BladeCenter chassis. The object is to ensure that there is no single point of failure which can cause the application(s) to become unavailable or unreachable.

The following are some important design considerations to try to ensure connectivity is maintained under various failure scenarios:

- For NIC teaming to work properly with trunk failover, you must have external Layer 2 connectivity between the GbESMs. This can be done by cabling the GbESM modules directly to each other or by connecting them both to the same collection of upstream switches.
- VRRP also requires a Layer 2 connection between switches. This connection must carry all the VLANs which have a VRRP instance configured.
- To provide robust HA in a Layer 3 design:
  - The two Nortel GbESMs should be configured with VRRP.
  - The blade servers need to be using the VRRP address(es) for the VLANs where they are configured as their default gateway.
  - It is possible to use VRRP (or equivalent) on the upstream switches as well to provide an even more robust HA design.

Note that the failure of a NIC within the blade server, the failure of a link between the GbESM and the blade server, and the hard failure of a GbESM would all result in a link down condition and would be successfully detected by NIC Teaming without the use of trunk failover.

7.4 Guidelines for attaching the BladeCenter to a network

This section contains information about things to consider when attaching the IBM @server BladeCenter to a network. We highly recommend that you review this entire section prior to any initial configuration changes. The topologies presented in this chapter discuss attaching the IBM @server BladeCenter to an external infrastructure.

7.4.1 Guidelines and comments

The following sections present comments and recommendations that are related to the various BladeCenter components which are used in the examples in this chapter.

Cable type selection (cross-over or straight-through)

Selection of the cable type (cross-over or straight-through) to use between the Nortel Networks L2/3 GbESM and an external switch is important. Although both a straight-through and a cross-over have been shown to work correctly in the lab during the creation of this document, there are certain times (such as when hard-coding link speed or duplex characteristics) when only a cross-over cable will work. Based on this, we strongly recommend that you use a cross-over cable between the Nortel Networks L2/3 GbESMs and upstream switches. This helps ensure that the link always works under all possible conditions.

Fiber connections

The fiber connectors on the Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter must be *Multi-Mode Fiber (MMF)*. Either 50 or 62.5 micron fiber can be used. Single-Mode (9 micron) fiber is not supported. The fiber should be terminated with LC connectors on both ends.

Speed or duplex selection

The decision to allow a port to negotiate its speed and duplex automatically or to force it to a set value is a subject of frequent debate. Testing in the lab has shown that the Nortel Networks L2/3 GbESM can negotiate the link correctly when attaching to external switches. In particular, with Gigabit connections, we strongly recommend that you use auto-negotiation.

> **Important:** Although you can attach the Nortel Networks Layer 2/3 Copper Gigabit Ethernet Switch Module for IBM @server BladeCenters to external switches at 10 or 100 Mb, in production environments we strongly recommend that you use 1 Gb connections. When using the Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter, you must use 1 Gb connections because that is the only speed that is supported.

7.5 Base configurations for examples in this document

Before discussing the specifics of each configuration, it is necessary to outline the hardware and software used during the experiments.

7.5.1 Hardware and software used for lab environment

It should be noted that the choice of the 3560Gs and their components was made based on the assumption that the IBM @server BladeCenter is being deployed in a mission-critical data center environment, where high availability and performance are of utmost importance. Also, the 3560G Cisco switch supports the standards that are used in each experiment for Layer 2 and 3 switching.

IBM @server BladeCenter configuration

The IBM @server BladeCenter was configured as follows:

- One BladeCenter chassis (8677-1XZ) with:
  - Four HS20 blades (8678-2ZZ) in slots 1 through 4
    - One 2.4 GHz CPU
    - One 40 GB hard disk
    - 2560 MB of memory
    - BIOS build BRE134AUS, version displayed is 1.09
    - Diagnostic build BRYT18AUS
    - Integrated System Management Processor (ISMP) build BR8T35A
    - Windows 2003 Standard Edition operating system
    - Broadcom firmware version 3.21
    - Broadcom driver version 8.22.1.0
    - BASP software version 8.1.4
  - One HS40 blade (8839-7HX) in slot 5
    - Four 3.0 GHz CPUs
    - One 40 GB hard disk
    - 2 GB of memory

      - BIOS build SBJT58AUS
      - Diagnostic build SBY113AUS
      - Integrated System Management Processor (ISMP) build BRMK27A
      - Windows 2003 Standard Edition operating system
      - Intel driver version 7.3.13.0

  - Four 2000 watt power supplies in the BladeCenter chassis
  - Two of the Nortel Networks Layer 2/3 Copper Gigabit Ethernet Switch Module for IBM @server BladeCenters (#26K6524)
  - One BladeCenter Management Module (#59P2960) with firmware version BRET79A

Nortel Networks L2/3 GbESM

Two Nortel Networks Layer 2/3 Copper Gigabit Ethernet Switch Module for IBM @server BladeCenter running code 1.0.1.6.

Cisco 3560G switch

Two Cisco Systems 3560G series PoE24 running IOS version 12.2 (25) SEB1.

Additional hardware used during the experiments

The following hardware was also used in the lab examples. The Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter was used in verifying configurations that were based off the copper GbESM. The only difference between both is the PHY layer from copper to fiber. The Extreme Networks switches were used to show examples of the Nortel GbESMs configured with a different switch vendor.

- Two Nortel Networks Layer 2/3 Fiber Gigabit Ethernet Switch Module for IBM @server BladeCenter (#26K6528) running code 1.0.1.6
- Two Extreme Networks Summit® 400-48t switches running ExtremeWare version 7.2e.1

7.5.2 Preconfiguration preparation

The configurations in this document were built off each other from a very basic topology to the more complex final solutions at Layer 3. Each example is a progression from the first. Configurations after each experiment were not reset to start from scratch. The basic topology section starts off with the Cisco and Nortel switches at default settings. Any configuration changes from there are documented.

Important: If working in a production network, be sure to understand the consequences of any commands that are issued. Failure to completely understand the operation of commands can lead to network down conditions.

Note: Available features and command syntax can be different with different versions of code. This document was prepared using the features and syntax from the aforementioned revisions of code, and as such, might vary from other revisions. For complete and current lists of available features and commands for these products, visit the IBM or Nortel Web sites.

7.5.3 Base configuration common to all examples

This section lists some established configuration options that are common to all of the examples. These are only for demonstration purposes in the examples and might or might not be duplicated in your particular environment.

Management Module settings for Nortel GbESMs

Each Nortel GbESM is configured with an IP address for the MGT1 ports (see Figure 5-2 on page 31 for more detail):

- GbESM_1 is configured with 9.42.171.243 and mask 255.255.255.0
- GbESM _2 is configured with 9.42.171.244 and mask 255.255.255.0
- The default gateway is set to the internal interface of the Management Module 9.42.171.242
- Both GbESMs have Fast POST enabled
- Both GbESMs have External Ports enabled
- Both GbESMs have External Management over all Ports enabled
- Both GbESMs have Preserve new IP configuration on all resets enabled

We do not discuss the Management Module configuration from this point forward. There are no changes to the settings listed above.

IP address and VLAN scheme

The IP address for all lab examples is written as 10.*x*.0.*y*. The *x* identifies the VLAN while the *y* identifies the switch. An example of this would be 10.99.0.245. The 99 identifies VLAN 99 which is a management VLAN created for the examples. The last octet being 245 distinguishes the IP address as one belonging to Core1. Core1 will always have the last octet of 245. The last octet for Core2 is 246, GbESM_1 is 243, and GbESM_2 is 244. Each blade will have the slot number for the last octet (example: slot 1 would be 1). The net mask is 255.255.255.0 throughout these examples. Also of note, GbESM_1 is the switch module in slot 1 of the BladeCenter chassis.

All example configurations have some combination of the following VLANs configured: VLAN 1, 5, 10, 20, 35, 36, 45, 46, or 99.

> **Note:** The VLANs chosen here are only for the purposes of demonstration and might or might not be a part of your particular network.

All configurations assume that VLANs carried on 802.1Q trunks are limited to only those that are necessary (this is good security practice). Additionally, VLAN 5 is created for use as the untagged or native VLAN on 802.1Q links.

Blade server configuration with BASP

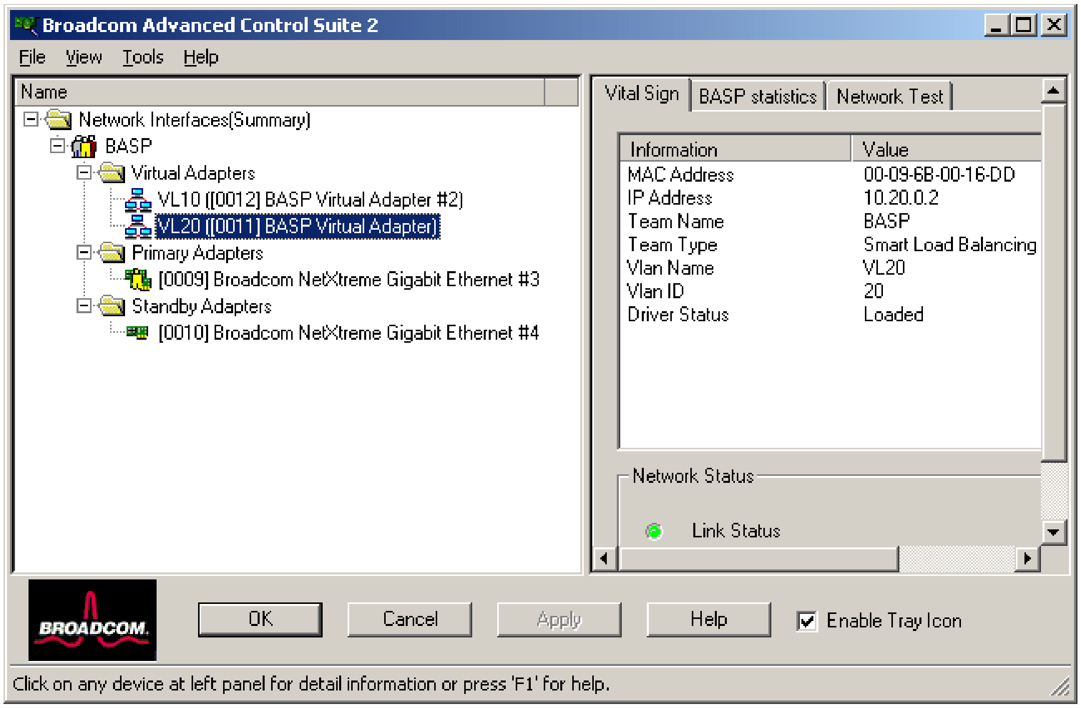
A team of both blade Ethernet interfaces is made by launching the BASP Advanced Control Suite software. Smart Load Balance and Failover is the teaming feature used in this document. The following VLANs are placed on the blade servers (exact number and placement depends on trunking and teaming for the given example):

- Blade Server 1: VLAN 20
- Blade Server 2: VLAN 10, 20
- Blade Server 3: VLAN 10, 20
- Blade Server 4: VLAN 99

For blade server 1, a team is made using the first Ethernet interface as the primary and the second as standby. The IP address for the new BASP interface is set to 10.20.0.1. This is called Active/Standby mode.

Note: First Ethernet interface refers to the blade's physical connection to the first Ethernet switch module in slot 1.

For blade server 2, a team is made the same way as blade server 1. Here two VLANs are created (VL10 and VL20). Both of these must be set as tagged. Each VLAN is given an IP address (VL10 = 10.10.0.2; VL20 = 10.20.0.2). Note here that the VL10 and VL20 are only names for the VLANs displayed in the BASP configuration window. The actual VLAN ID must be set to the same ID the switches are carrying. This means that VL10 must have a VLAN ID of 10, and VL20 must have its VLAN ID set to 20 according to the scheme being used in this document. Figure 7-2 shows the BASP configuration for blade server 2.



Name

- Network Interfaces(Summary)
  - BASP
    - Virtual Adapters
      - VL10 ([0012] BASP Virtual Adapter #2)
      - VL20 ([0011] BASP Virtual Adapter)
    - Primary Adapters
      - [0009] Broadcom NetXtreme Gigabit Ethernet #3
    - Standby Adapters
      - [0010] Broadcom NetXtreme Gigabit Ethernet #4

Information	Value
MAC Address	00-09-6B-00-16-DD
IP Address	10.20.0.2
Team Name	BASP
Team Type	Smart Load Balancing
Vlan Name	VL20
Vlan ID	20
Driver Status	Loaded

Network Status

● Link Status

☑ Enable Tray Icon

Click on any device at left panel for detail information or press 'F1' for help.

Figure 7-2 Blade server 2 BASP configuration

Blade server 3 is configured so that the first Ethernet interface has its IP address on VLAN 20. The second Ethernet interface of the blade is on VLAN 10. BASP is not configured on this blade. Only Windows networking was used to configure the IP addresses.

Blade server 4 is configured much the same way as blade 1. However, this blade only uses VLAN 99. The IP address is set to 10.99.0.4 to keep with the same IP scheme. This allows for a blade server inside the chassis to be on the management VLAN.

7.6 Basic Layer 2 entry topology

This section lays out the first configuration from which each continuing section builds. This topology uses 802.1Q tagging for carrying several different networks. The topology also includes static port aggregation (Etherchannel) with trunk failover. Something that should be noted is that this offering is basic and has limited redundancy that relies on port aggregation and trunk failover.

7.6.1 Layer 2 configuration with 802.1Q tagging and trunk failover

Figure 7-3 illustrates the first basic Layer 2 topology.

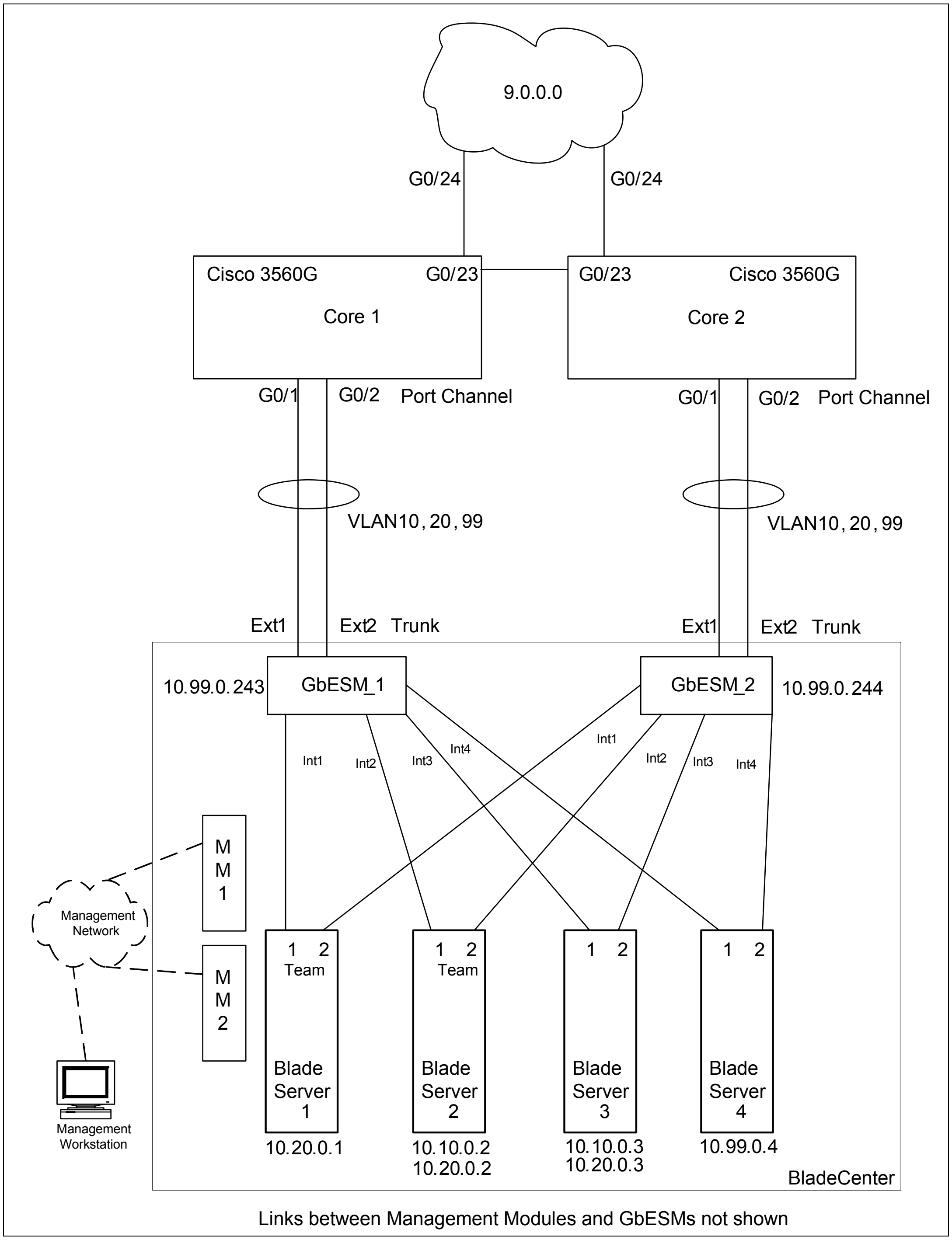


Figure 7-3 Basic Layer 2 topology with 802.1Q tagging and trunk failover

Summary of disconnect procedure to be performed for each example

When performing initial configurations or making changes to existing configurations that might have an impact on Spanning Tree (such as changing link aggregation), it is recommended that you leave connections uncabled or shut down prior to making the configuration changes. This reduces the likelihood of any temporary Spanning Tree loops and possible network-down conditions that might result in the process of adding or changing configurations.

- Shut down the ports on GbESM_1 and GbESM_2

```
/oper/port EXT1/dis
/oper/port EXT2/dis
```

- Shut down the ports on Core1 and Core2

```
conf t
int range g0/1-2
shut
```

Cisco 3560G switch configuration

In this topology, each Cisco switch connects two ports (g0/1 and g0/2) to its adjoining Nortel GbESM (EXT1 and EXT2). These aggregated links are to carry VLAN 5, 10, 20, and 99. VLAN 99 is meant to be a management VLAN, and the only blade server in the chassis to have access to it is blade server 4. The native VLAN here is 5.

In the switch configuration mode, create the VLANs and IP interfaces first, as shown in Example 7-1.

Example 7-1 Create the VLANs and IP interfaces first

```
:
conf t
!
vlan 5
      name native
vlan 10
      name vlan_green
vlan 20
      name vlan_red
vlan 99
      name mgmt
```

By issuing **do sh vlan**, you can see the intended VLAN layout, as shown in Example 7-2.

Example 7-2 Verifying the VLAN setup

```
Core1(config)#do sh vlan

VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active    Gi0/3, Gi0/4, Gi0/5, Gi0/6
                                                Gi0/7, Gi0/8, Gi0/9, Gi0/10
                                                Gi0/11, Gi0/12, Gi0/13, Gi0/14
                                                Gi0/15, Gi0/16, Gi0/17, Gi0/18
                                                Gi0/19, Gi0/20, Gi0/21, Gi0/22
                                                Gi0/25, Gi0/26, Gi0/27, Gi0/28
5    native                           active
10   vlan_green                       active
20   vlan_red                         active
99   mgmt                             active
1002 fddi-default                     act/unsup
```

```
1003 token-ring-default               act/unsup
1004 fddinet-default                  act/unsup
1005 trnet-default                    act/unsup

VLAN Type  SAID       MTU   Parent RingNo BridgeNo Stp  BrdgMode Trans1 Trans2
---- ----- ---------- ----- ------ ------ -------- ---- -------- ------ ------
1    enet  100001     1500  -      -      -        -    -        0      0
5    enet  100005     1500  -      -      -        -    -        0      0
10   enet  100010     1500  -      -      -        -    -        0      0
20   enet  100020     1500  -      -      -        -    -        0      0

VLAN Type  SAID       MTU   Parent RingNo BridgeNo Stp  BrdgMode Trans1 Trans2
---- ----- ---------- ----- ------ ------ -------- ---- -------- ------ ------
99   enet  100099     1500  -      -      -        -    -        0      0
1002 fddi  101002     1500  -      -      -        -    -        0      0
1003 tr    101003     1500  -      -      -        -    -        0      0
1004 fdnet 101004     1500  -      -      -        ieee -        0      0
1005 trnet 101005     1500  -      -      -        ibm  -        0      0

Remote SPAN VLANs
------------------------------------------------------------------------------


Primary Secondary Type              Ports
------- --------- ----------------- ------------------------------------------
```

Note: If there are any unwanted VLANs, you can remove them with the `no vlan #` command (where # is the number of the VLAN to remove). The VLAN information is not included when dumping the switch configuration with the `sh run` command. VLAN information is stored separately in a vlan.dat file.

Example 7-3 creates IP interfaces to be used later.

Example 7-3 Create IP addresses

```
interface Vlan10
      ip address 10.10.0.245 255.255.255.0         ! ip address 10.10.0.246 255.255.255.0
interface Vlan20
      ip address 10.20.0.245 255.255.255.0         ! ip address 10.20.0.246 255.255.255.0
interface Vlan99
      ip address 10.99.0.245 255.255.255.0         ! ip address 10.99.0.246 255.255.255.0
```

Important: When looking at the configuration commands presented, commands toward the left margin are for Core1 or GbESM_1 switches. Where the GbESM_2 or Core2 switch configurations differ, the commands presented within comment syntax (/* for GbESM, ! for Cisco, # for Extreme) are for the counterpart switch of the same type, GbESM_2 or Core2.

To start setting up the ports:

1. Access the interface level for g0/1 and g0/2 with the following:

   ```
   interface range g0/1-2
   ```

2. Enable 802.1Q tagging with VLAN 5 untagged and allow the VLANs which should be carried over the aggregated ports:

   ```
   switchport trunk encapsulation dot1q
   switchport trunk native vlan 5
   switchport trunk allowed vlan 5,10,20,99
   switchport mode trunk
   ```

3. Disable the Cisco proprietary dynamic trunk protocol DTP:

   ```
   switchport nonegotiate
   ```

4. Enable aggregation by choosing a channel-group number and mode on:

   ```
   channel-group 1 mode on
   ```

This creates and enables a virtual interface called Port-channel1 (po1). Remember that the ports are still disabled.

Now, to simulate the core network, Example 7-4 links both Cisco switches together (g0/23) and also links them further upstream (g0/24):

Example 7-4 Links

```
interface g0/23
      switchport trunk encapsulation dot1q
      switchport trunk native vlan 5
      switchport trunk allowed vlan 5,10,20,99
      switchport mode trunk
!
interface g0/24
      no switchport
      ip address 9.42.171.245 255.255.255.0       ! ip address 9.42.171.246 255.255.255.0
!
ip default-gateway 9.42.171.3
```

The host name was set in order to easily identify the switch on the command line. All configuration changes were then saved:

```
hostname Core1                                     /* hostname Core2
end
wri mem
```

All other configuration to the switch can be left as default. The above examples set up aggregation for ports g0/1 and g0/2 and allow for the channel to carry the VLANs used in the experiment. The Cisco switch itself can also be accessed on the management VLAN 99.

Nortel Networks L2/3 GbESM switch configuration

In Example 7-8 on page 76, each static trunk created on the Cisco switch connects to EXT1 and EXT2 of its adjoining Nortel GbESM. Again, as with the Cisco, the native VLAN here is VLAN 1. This is also default for the Nortel switch.

First, ports EXT1 and EXT2 must be configured as tagged:

1. Tagging on the INT ports is enabled by default due to all INT ports being members of VLAN 1 and 4095:

   ```
   /c/port EXT1
         tag ena
   /c/port EXT2
         tag ena
   ```

2. Create the VLANs, enable them, and add ports, as shown in

Example 7-5 Create and enable VLANs

```
   /c/l2/vlan 5
      ena
      name “Native”
      def EXT1 EXT2
/c/l2/vlan 10
      ena
      name "VLAN_Green"
      def INT2 INT3 EXT1 EXT2                        /* def INT2 EXT1 EXT2
/c/l2/vlan 20
      ena
      name "VLAN_Red"
      def INT1 INT2 EXT1 EXT2                        /* def INT1 INT2 INT3 EXT1 EXT2
/c/l2/vlan 99
      ena
      name "MGMT"
      def INT4 EXT1 EXT2
```

3. In Example 7-6, INT ports must have PVIDs set. Blade servers on INT1, INT3, and INT4 as configured are untagged. The external ports should have PVID set to 5.

Example 7-6 PVIDs set

```
/c/port EXT1
      pvid 5
/c/port EXT2
      pvid 5
/c/port INT1
      pvid 20
/c/port INT3
      pvid 10                                        /* pvid 20
/c/port INT4
      pvid 99
```

There is one change on GbESM_2 for INT3. This is so that blade server 3, connecting through INT3, on GbESM_1 will be on VLAN 10. It will be on VLAN 20 through GbESM_2. INT2 does not need a PVID set on either Nortel because by default it is already set to 1. The BASP setting for both VL10 and VL20 is set to tagged VLAN.

Remove EXT1 and EXT2 from the default VLAN 1 (as a security precaution):

```
/c/l2/vlan 1
      rem EXT1
      rem EXT2
```

Unlike the Cisco switch, VLAN information is included when dumping the switch configuration with the **/c/d** command. VLAN information can be displayed with the **/i/l2/vlan** command.

Example 7-7 shows what has been configured so far.

Example 7-7 Verifying the VLAN setup of the Nortel GbESM

```
>> Main# /i/l2/vlan
VLAN                Name               Status Ports
----  -------------------------------- ------ -----------------------
1     Default VLAN                       ena   INT1-INT14 EXT3-EXT6
5     Native                             ena   EXT1 EXT2
10    VLAN_Green                         ena   INT2 EXT1 EXT2
20    VLAN_Red                           ena   INT1-INT3 EXT1 EXT2
99    MGMT                               ena   INT4 EXT1 EXT2
4095  Mgmt VLAN                          ena   INT1-MGT2
```

Spanning tree can be disabled for this topology. Because all VLANs are members of stg 1, by default, the following turns off the group:

```
/c/l2/stg 1 /off
```

To set up the trunk between EXT1 and EXT2 to interface with each Cisco switch, use the following (also, trunk failover is enabled with the syntax **/c/l2/trunk 1/fail ena**):

```
/c/l2/trunk 1
      add ext1
      add ext2
      ena
      fail ena
```

The Nortel GbESMs can be managed on VLAN 99:

```
/c/l3/if 99
      ena
      addr 10.99.0.243                                /* addr 10.99.0.244
      mask 255.255.255.0
      vlan 99
```

For passing management traffic through the network, each Nortel GbESM should reference the upstream Cisco switches as gateways on VLAN 99:

```
/c/l3/gw 1
      ena
      addr 10.99.0.245                                /* addr 10.99.0.246
```

It is important here to enable the gateway with the **/c/l3/gw 1/ena** command or the entry remains disabled.

Additionally, the sysName was set in order to easily identify the switch. All configuration changes were applied and then saved to the flash:

```
/c/sys/ssnmp
      name “GbESM_1”                                  /* name “GbESM_2”

apply
save
```

Re-enable the ports

After verifying the correct cabling between all the devices, the ports can be re-enabled.

- Enabling the ports on GbESM_1 and GbESM_2

```
/oper/port EXT1/ena
/oper/port EXT2/ena
```

- Enabling the ports on Core1 and Core2

```
conf t
int range g0/1-2
no shut
```

After the ports have been enabled, the aggregated links should be functioning correctly. Example 7-8 shows the trunk in forwarding state on the Nortel GbESM.

Example 7-8 Verifying trunk link is in forwarding state on the Nortel GbESM

```
>> Layer 2# /i/l2/trunk
Trunk group 1: Enabled
failover ena, port state:
 EXT1: STG  1 forwarding
 EXT2: STG  1 forwarding
```

Complete configuration snapshots

Complete configuration files are shown for the GbESM (Example 7-9) and upstream Core switches (Example 7-10 on page 77). One configuration for each type of switch is shown, with notes where the second switch of the same type has a different configuration.

Example 7-9 Basic topology GbESM configuration - tagging and trunk failover

```
>> Main# /c/d
script start "Layer 2-3 Gigabit Ethernet Switch Module for IBM eServer BladeCent
er" 4  /**** DO NOT EDIT THIS LINE!
/* Configuration dump taken 13:44:45 Wed Jun 22, 2005
/* Version 1.0.1.6,  Base MAC address 00:11:f9:36:b7:00
/* GbESM_1
/c/sys/ssnmp
        name "GbESM_1"                              /* name “GbESM_2”
/c/port INT1
        pvid 20
/c/port INT3
        pvid 10                                     /* pvid 20
/c/port INT4
        pvid 99
/c/port EXT1
        tag ena
        pvid 5
/c/port EXT2
        tag ena
        pvid 5
/c/l2/vlan 1
        def INT1 INT2 INT3 INT4 INT5 INT6 INT7 INT8 INT9 INT10 INT11 INT12 INT13
 INT14 EXT3 EXT4 EXT5 EXT6
/c/l2/vlan 5
        ena
        name "Native"
        def EXT1 EXT2
/c/l2/vlan 10
        ena
        name "VLAN_Green"
        def INT2 INT3 EXT1 EXT2                    /* def INT2 EXT1 EXT2
/c/l2/vlan 20
        ena
        name "VLAN_Red"
        def INT1 INT2 EXT1 EXT2                    /* def INT1 INT2 INT3 EXT1 EXT2
/c/l2/vlan 99
        ena
        name "MGMT"
        def INT4 EXT1 EXT2
/c/l2/stg 1/off
/c/l2/stg 1/clear
/c/l2/stg 1/add 1 5 10 20 99
/c/l2/trunk 1
        ena
        failovr ena
        add EXT1
        add EXT2
/c/l3/if 99
        ena
        addr 10.99.0.243                           /* addr 10.99.0.244
        mask 255.255.255.0
        broad 10.99.0.255
        vlan 99
/c/l3/gw 1
        ena
        addr 10.99.0.245                           /* addr 10.99.0.246
/
script end  /**** DO NOT EDIT THIS LINE!
```

Example 7-10 Basic topology Core switch configuration - tagging and etherchannel

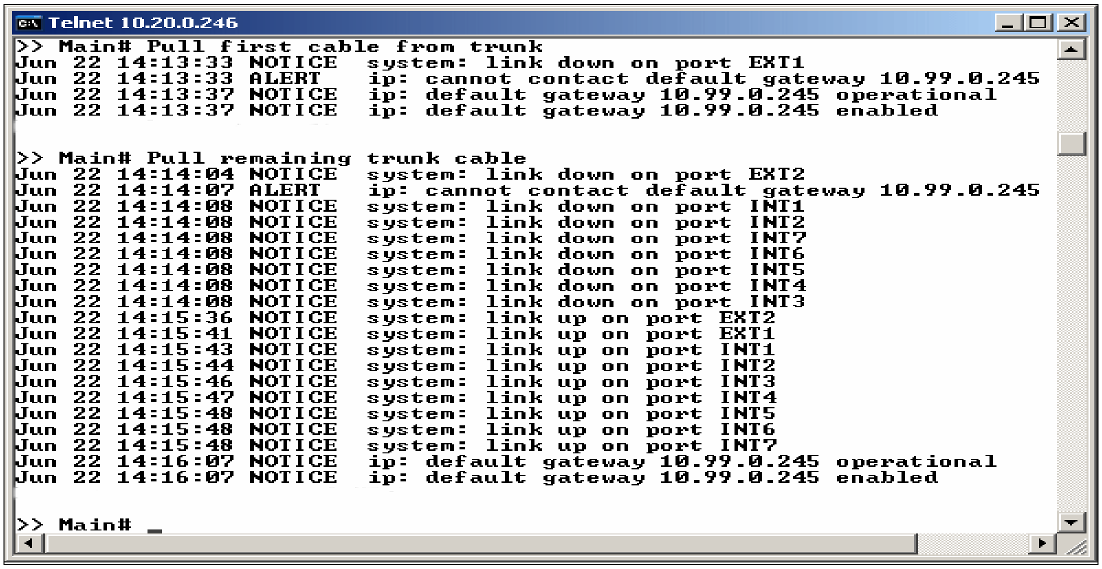
```
Core1#sh run
Building configuration...

Current configuration : 2383 bytes
!
version 12.2
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname Core1                                      ! hostname Core2
!
enable password cisco
!
no aaa new-model
ip subnet-zero
no ip domain-lookup
!
!
!
!
!
!
no file verify auto
spanning-tree mode pvst
```

```
spanning-tree extend system-id
!
vlan internal allocation policy ascending
!
!
interface Port-channel1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
interface GigabitEthernet0/1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 1 mode on
!
interface GigabitEthernet0/2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 1 mode on
!
interface GigabitEthernet0/3
!
interface GigabitEthernet0/4
!
interface GigabitEthernet0/5
!
interface GigabitEthernet0/6
!
interface GigabitEthernet0/7
!
interface GigabitEthernet0/8
!
interface GigabitEthernet0/9
!
interface GigabitEthernet0/10
!
interface GigabitEthernet0/11
!
interface GigabitEthernet0/12
!
interface GigabitEthernet0/13
!
interface GigabitEthernet0/14
!
interface GigabitEthernet0/15
!
interface GigabitEthernet0/16
!
interface GigabitEthernet0/17
!
interface GigabitEthernet0/18
!
```

```
interface GigabitEthernet0/19
!
interface GigabitEthernet0/20
!
interface GigabitEthernet0/21
!
interface GigabitEthernet0/22
!
interface GigabitEthernet0/23
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
!
interface GigabitEthernet0/24
 no switchport
 ip address 9.42.171.245 255.255.255.0             ! ip address 9.42.171.246 255.255.255.0
!
interface GigabitEthernet0/25
!
interface GigabitEthernet0/26
!
interface GigabitEthernet0/27
!
interface GigabitEthernet0/28
!
interface Vlan1
 no ip address
 shutdown
!
interface Vlan10
 ip address 10.10.0.245 255.255.255.0              ! ip address 10.10.0.246 255.255.255.0
!
interface Vlan20
 ip address 10.20.0.245 255.255.255.0              ! ip address 10.20.0.246 255.255.255.0
!
interface Vlan99
 ip address 10.99.0.245 255.255.255.0              ! ip address 10.99.0.246 255.255.255.0
!
ip default-gateway 9.42.171.3
ip classless
ip http server
ip http secure-server
!
!
!
control-plane
!
!
line con 0
line vty 0 4
 password cisco
 no login
line vty 5 15
 no login
!
!
end
```

7.6.2 Basic topology conclusions

Testing for this configuration involved pulling cables from the g0/1 and g0/2 ports of Core1 while pings were running. Removing the cables one by one shows that first the trunk redundancy switches the link. After removing the second cable, the trunk failover brings down the INT ports of GbESM_1. If the Nortel switch is being monitored, Figure 7-4 illustrates what you see. After 1 minute and 30 seconds of pulling the second cable of the trunk, we reattached both cables to Core1.



```
>> Main# Pull first cable from trunk
Jun 22 14:13:33 NOTICE  system: link down on port EXT1
Jun 22 14:13:33 ALERT   ip: cannot contact default gateway 10.99.0.245
Jun 22 14:13:37 NOTICE  ip: default gateway 10.99.0.245 operational
Jun 22 14:13:37 NOTICE  ip: default gateway 10.99.0.245 enabled

>> Main# Pull remaining trunk cable
Jun 22 14:14:04 NOTICE  system: link down on port EXT2
Jun 22 14:14:07 ALERT   ip: cannot contact default gateway 10.99.0.245
Jun 22 14:14:08 NOTICE  system: link down on port INT1
Jun 22 14:14:08 NOTICE  system: link down on port INT2
Jun 22 14:14:08 NOTICE  system: link down on port INT7
Jun 22 14:14:08 NOTICE  system: link down on port INT6
Jun 22 14:14:08 NOTICE  system: link down on port INT5
Jun 22 14:14:08 NOTICE  system: link down on port INT4
Jun 22 14:14:08 NOTICE  system: link down on port INT3
Jun 22 14:15:36 NOTICE  system: link up on port EXT2
Jun 22 14:15:41 NOTICE  system: link up on port EXT1
Jun 22 14:15:43 NOTICE  system: link up on port INT1
Jun 22 14:15:44 NOTICE  system: link up on port INT2
Jun 22 14:15:46 NOTICE  system: link up on port INT3
Jun 22 14:15:47 NOTICE  system: link up on port INT4
Jun 22 14:15:48 NOTICE  system: link up on port INT5
Jun 22 14:15:48 NOTICE  system: link up on port INT6
Jun 22 14:15:48 NOTICE  system: link up on port INT7
Jun 22 14:16:07 NOTICE  ip: default gateway 10.99.0.245 operational
Jun 22 14:16:07 NOTICE  ip: default gateway 10.99.0.245 enabled

>> Main# _
```

Figure 7-4 INT ports blocked during a trunk failover test

We also noticed in testing this configuration that there was no difference in the failover behavior between LACP and a static configured trunk. LACP is not shown as configured in the examples because the current generally available code release does not support trunk failover with LACP.

> **Note:** LACP with trunk failover is a feature to be included in an upcoming release. During this experiment a early version of this code was tested and the feature does work.

One ping was lost on the first cable pull, while only three were lost on the second when the failover occurred. At most, only three pings were lost when the cables were reattached and failback occurred.

7.7 Advanced Layer 2 topology sample configurations

This example is an extension to the basic configuration described in 7.6, “Basic Layer 2 entry topology” on page 69, because cross connections between the GbESMs and the upstream switches are established now. See Figure 7-5 on page 81 for the topology used in this example. The crosslinks increase the redundancy and provide more flexibility regarding the configuration of the BladeCenter components (for example, trunk failover is not needed for full redundancy) but it costs the need for loop prevention. This is commonly reached through the usage of a Spanning Tree Protocol, blocking dedicated ports to break the loop on Layer 2. Another approach to break the loop, based on the Layer 3 capabilities of the GbESMs, is shown in 7.8, “Layer 3 topology sample configurations” on page 108. The initial configuration for all the switches in the following examples are the same as those in 7.6, “Basic Layer 2 entry topology” on page 69, except that IEEE 802.ad dynamic port aggregation (LACP) is used on all the upstream connections rather than the static Etherchannel.

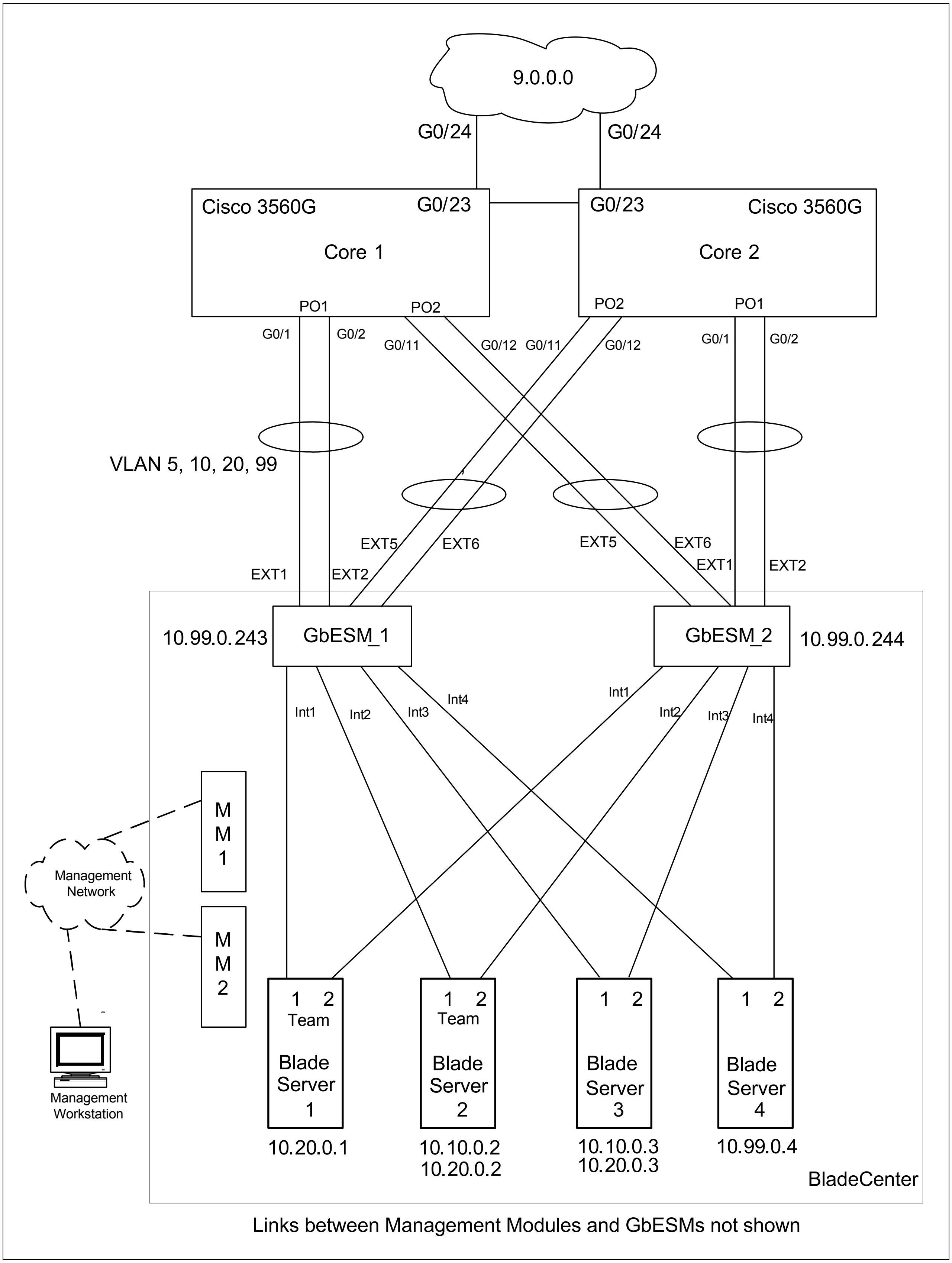


Figure 7-5 Advanced Layer 2 topology

The Advanced Layer 2 topology offers a good compromise between performance and high availability. It is made up of dual GbESMs, each with two, 2-port aggregated links, going to a pair of Core switches, each Core switch joined to each other via a single link (simulating a Layer 2 network beyond the switches). The 2-port aggregation itself provides for higher performance and the second 2-port aggregation link provides for full redundancy on a connection-loss or a switch-failure situation.

In the examples presented in this chapter, the Core switches are always forced to be the Spanning Tree Root Bridges. Having the Root Bridge directly attached to the GbESM is not necessarily recommended in redundant configurations, as flow patterns can become less than obvious. But it is highly recommended to have the Root Bridge outside the BladeCenter.

> **Important:** There is a high probability that any existing network will already have a desired switch configured as the root. It is very important that you understand the proper selection of the root bridge and that the GbESM not be allowed to become the root bridge. Allowing the GbESM to become the root bridge can result in sub-optimal data flow within the Layer 2 network.

Summary of disconnect procedure, to be performed for each example

When performing initial configurations or making changes to existing configurations that might have an impact on Spanning Tree (such as changing link aggregation), it is recommended that you leave connections un-cabled, or shut down, prior to making the configuration changes. This will reduce the likelihood of any temporary Spanning Tree loops and possible network-down conditions that might result in the process of adding or changing configurations.

- Shut down the ports on GbESM_1 and GbESM_2

```
/oper/port EXT1/dis
/oper/port EXT2/dis
/oper/port EXT5/dis
/oper/port EXT6/dis
```

- Shut down the ports on Core1 and Core2

```
conf t
int range g0/1-2, g0/11-12
shut
```

7.7.1 Dynamic link aggregation IEEE 802.3ad (LACP)

In contrast to the static trunks configured in 7.6, “Basic Layer 2 entry topology” on page 69, dynamic Link Aggregation Control Protocol (LACP) is used now for the connections GbESM_1 to Core1 and GbESM_2 to Core2. The additional cross connections, GbESM_1 ports EXT5-6 to ports G0/11-12 on Core2 and similar GbESM_2 ports EXT5-6 to ports G0/11-12 on Core1 are configured as LACP trunks as well, including IEEE 802.1Q tagging (VLAN 5 untagged/native). Remember to shut down the ports before starting the configuration changes as mentioned before.

Configure the Nortel GbESMs for IEEE 802.1Q tagged LACP trunk

Enable IEEE 802.1Q tagging on ports EXT5-6 and set the PVID to the VLAN which should not be tagged (must be equal to the native VLAN on the Cisco switches):

```
/c/port EXT5
      tag ena
      pvid 5
/c/port EXT6
      tag ena
      pvid 5
```

Add ports EXT5-6 to the VLANs that should be carried through the LACP trunks:

```
/c/l2/vlan 5
      add EXT5
      add EXT6
/c/l2/vlan 10
      add EXT5
      add EXT6
/c/l2/vlan 20
      add EXT5
      add EXT6
/c/l2/vlan 99
      add EXT5
      add EXT6
```

Remove EXT5-6 from the default VLAN 1 (as a security precaution):

```
/c/l2/vlan 1
      rem EXT5
      rem EXT6
```

Note: The same should have been done already for ports EXT1 and EXT2 (see 7.6.1, “Layer 2 configuration with 802.1Q tagging and trunk failover” on page 70). If not, the above steps must be repeated with EXT5, EXT6 replaced by EXT1, EXT2.

Configure LACP to aggregate ports EXT1, EXT2 and EXT5, EXT6 each by choosing the mode `active` (if desired `standby` would be another option) for the lower port of the pair:

```
/c/l2/lacp/port EXT1
      mode active
/c/l2/lacp/port EXT5
      mode active
```

Look for the `admin key` used on the lower ports (EXT1 and EXT5) with **`/i/l2/lacp/dump`** and set it accordingly on the higher ports EXT2 and EXT6:

```
/c/l2/lacp/port EXT2
      mode active
      adminkey 17
/c/l2/lacp/port EXT6
      mode active
      adminkey 21
```

Because this configuration is based on the one in 7.6, “Basic Layer 2 entry topology” on page 69, the static port aggregation for ports EXT1 and EXT2 must be disabled with the **`/c/trunk 1/dis`** command. Finally `apply` must be entered to execute the configuration changes. Remember that the ports EXT1, EXT2, EXT5, and EXT6 are still disabled.

Configure the Cisco Core1 and Core2 for tagged LACP trunks

In the configuration mode, access the interface level for G0/11 and G0/12:

```
conf t
interface range G0/11-12
```

Enable IEEE 802.1Q tagging with VLAN 5 untagged and allow only the VLANs which should be carried over the LACP trunk:

```
switchport trunk encapsulation dot1q
switchport trunk native vlan 5
switchport trunk allowed vlan 5,10,20,99
switchport mode trunk
```

Disable the Cisco proprietary dynamic trunk protocol DTP:

```
switchport nonegotiate
```

Enable LACP by choosing a channel-group number and use active mode:

```
channel-group 2 mode active
```

That creates virtual interface called Port-channel2 (or short Po2).

For the port range G0/1-2, there should be already a static trunk configured (see 7.6.1, “Layer 2 configuration with 802.1Q tagging and trunk failover” on page 70). If not, the above steps must be repeated with G0/11-12 replaced by G0/1-2. Consequently only the aggregation mode must be changed for ports G0/1 and G0/2 (could be done for the virtual interface Po1 instead):

```
interface range G0/1-2
channel-group 1 mode active
end
```

Remember that the ports are still shut down.

Verify the LACP trunk status

Before enabling the configured ports to get the LACP trunk up, Spanning Tree must be activated. After that and re-enabling the ports, the status of the LACP trunks can be checked with:

- **`/i/l2/lacp/dump`** on the GbESM
- **`show lacp int`** and **`show lacp nei`** on the Cisco switch

Important: Do not enable the configured ports at this time to avoid a Layer 2 loop, what could be deadly for the whole network. Spanning Tree has to be configured and activated before. See 7.7.2, “Common Spanning Tree configuration - IEEE 802.1D and PVST” on page 84 and 7.7.3, “Rapid Spanning Tree IEEE 802.1w” on page 95.

7.7.2 Common Spanning Tree configuration - IEEE 802.1D and PVST

The GbESM interoperates with Cisco Per VLAN Spanning Tree (PVST) protocol using up to 16 Spanning Tree groups. This section shows how to configure separate Spanning Tree Groups (STG) for each VLAN using Cisco’s proprietary PVST and IEEE 802.1D (Common/Mono Spanning Tree).

The advantages of this approach are:

- Traffic can be distributed across several ports by blocking different ports on different VLANs. This is done by setting different root bridge priorities or port costs for different VLANs
- Some known issues which arise when connecting PVST and plain Mono Spanning Tree can be avoided. For more information, ssee *IBM @server BladeCenter Layer 2-7 Network Switching*, REDP-3755, at:

  http://www.redbooks.ibm.com/redpapers/pdfs/redp3755.pdf

Configure the Nortel GbESM_1 and GbESM_2 for Spanning Tree Groups

A current restriction of the GbESM is that when multiple Spanning Tree Groups are used, the Layer 3 forwarding function must be disabled:

```
/c/l3/frwd/off
```

The next step is to create a Spanning Tree bridge group for every active VLAN. Group 1 is fixed for VLAN 1, so we start with Group 2 for VLAN 5, Group 3 for VLAN 10, and so on. The bridge priority is set to the maximum value, so that it has the least priority in the Spanning Tree Root election process:

```
/c/l2/stg 2/clear
/c/l2/stg 2/brg/prior 65535
/c/l2/stg 2/add 5

/c/l2/stg 3/clear
/c/l2/stg 3/brg/prior 65535
/c/l2/stg 3/add 10

/c/l2/stg 4/clear
/c/l2/stg 4/brg/prior 65535
/c/l2/stg 4/add 20

/c/l2/stg 5/clear
/c/l2/stg 5/brg/prior 65535
/c/l2/stg 5/add 99
```

Finally, for our example, the Spanning Tree Group 1, controlling all other VLANs (in our case only VLAN 1), is disabled on the external ports, since we only want to carry VLANs 5,10, 20, and 99 on the earlier configured trunks:

```
/c/l2/stg 1/port EXT1/off
/c/l2/stg 1/port EXT2/off
/c/l2/stg 1/port EXT5/off
/c/l2/stg 1/port EXT6/off
```

Enter `apply` to execute the configuration changes.

> **Note:** The ports of the GbESM are enabled automatically after applying STG configuration changes. If necessary disable the ports afterwards with the **`/oper/port`** command as shown at the beginning of this section.

Configure the Cisco Core1 and Core2 for Per VLAN Spanning Tree

Cisco's proprietary PVST is the default value for Spanning Tree mode, so there is not necessarily something to configure. However, if you want to influence the Spanning Tree topology, you must modify the configuration. In this configuration, we decided to share the Root Bridge functionality for the different VLANs between the Core switches.

Ensure that PVST is selected as Spanning Tree mode, enter:

```
conf t
spanning-tree mode pvst
```

Core1 should become the Root Bridge for VLAN 5 and 10 while being backup Root for VLAN 20 and 99:

```
spanning-tree vlan 1-10 root primary
spanning-tree vlan 11-4094 root secondary
end
```

Vice versa for Core2, so that it will be elected as Root Bridge for VLAN 20 and 99, backing up the Root for VLAN 5 and 10:

```
spanning-tree vlan 1-10 root secondary
spanning-tree vlan 11-4094 root primary
end
```

Reenable the ports

After verifying the correct cabling between all the devices, the ports can be re-enabled.

- Enabling the ports on GbESM_1 and GbESM_2

```
/oper/port EXT1/ena
/oper/port EXT2/ena
/oper/port EXT5/ena
/oper/port EXT6/ena
```

- Enabling the ports on Core1 and Core2

```
conf t
int range g0/1-2, g0/11-12
no shut
```

Verify the Spanning Tree status

As shown in Example 7-11, the port status and the VLAN assignment could be verified by entering **info/link** and **info/port** on the GbESM.

Example 7-11 Verifying port status and VLAN assignment

```
>> GbESM_1 - Configuration# /i/link
------------------------------------------------------------------
Alias   Port   Speed    Duplex     Flow Ctrl      Link
-----   ----   -----   --------  --TX-----RX--   ------
INT1       1    1000     full     yes    yes       up
INT2       2    1000     full     yes    yes       up
INT3       3    1000     full     yes    yes       up
INT4       4    1000     full     yes    yes       up
INT5       5    1000     full     yes    yes       up
INT6       6    1000     full     yes    yes       up
INT7       7    1000     full     yes    yes       up
INT8       8    1000     full     yes    yes      down
INT9       9    1000     full     yes    yes      down
INT10     10    1000     full     yes    yes      down
INT11     11    1000     full     yes    yes      down
INT12     12    1000     full     yes    yes      down
INT13     13    1000     full     yes    yes      down
INT14     14    1000     full     yes    yes      down
MGT1      15     100     full     yes    yes       up
MGT2      16     100     full     yes    yes    disabled
EXT1      17    1000     full      no     no       up
EXT2      18    1000     full      no     no       up
```

```
EXT3     19      any      any     yes    yes    disabled
EXT4     20      any      any     yes    yes      down
EXT5     21     1000     full      no     no       up
EXT6     22     1000     full      no     no       up

>> GbESM_1 - Information# /i/port
Alias  Port  Tag  FAST  PVID      NAME              VLAN(s)
-----  ----  ---  ----  ----  -------------  -----------------------
INT1     1    y    n     20  INT1              1 4095   20
INT2     2    y    n      1  INT2              1 4095   10   20
INT3     3    y    n     10  INT3              1 4095   10
INT4     4    y    n     99  INT4              1 4095   99
INT5     5    y    n      1  INT5              1 4095
INT6     6    y    n      1  INT6              1 4095
INT7     7    y    n      1  INT7              1 4095
INT8     8    y    n      1  INT8              1 4095
INT9     9    y    n      1  INT9              1 4095
INT10   10    y    n      1  INT10             1 4095
INT11   11    y    n      1  INT11             1 4095
INT12   12    y    n      1  INT12             1 4095
INT13   13    y    n      1  INT13             1 4095
INT14   14    y    n      1  INT14             1 4095
MGT1    15    y    n   4095  MGT1           4095
MGT2    16    y    n   4095  MGT2           4095
EXT1    17    y    n      5  EXT1              5   10   20   99
EXT2    18    y    n      5  EXT2              5   10   20   99
EXT3    19    n    n      1  EXT3              1
EXT4    20    n    n      1  EXT4              1
EXT5    21    y    n      5  EXT5              5   10   20   99
EXT6    22    y    n      5  EXT6              5   10   20   99
```

The **info/l2/stg** command shows the Spanning Tree status for every active Spanning Tree Group. Verify that ports EXT5-6 of GbESM1 and EXT1-2 of GbESM2 are in BLOCKING status for the VLANs where Core1 is the Root Bridge, as shown in Example 7-12. Accordingly ports EXT1-2 of GbESM1 and EXT5-6 of GbESM2 should be in BLOCKING status for the VLANs where Core2 is the Root Bridge.

Example 7-12 Verifying the Spanning Tree status

```
>> GbESM_1 - Information# /i/l2/stg

------------------------------------------------------------------
Spanning Tree Group 1: Off (STP/PVST), FDB aging timer 300
VLANs:  1

Port  Priority   Cost      State       Designated Bridge     Des Port
----  --------   ----    ----------  ----------------------  --------
INT1        0      0     FORWARDING *
INT2        0      0     FORWARDING *
INT3        0      0     FORWARDING *
INT4        0      0     FORWARDING *
INT5        0      0     FORWARDING *
INT6        0      0     FORWARDING *
INT7        0      0     FORWARDING *
INT8        0      0      DISABLED  *
INT9        0      0      DISABLED  *
INT10       0      0      DISABLED  *
INT11       0      0      DISABLED  *
INT12       0      0      DISABLED  *
INT13       0      0      DISABLED  *
```

```
INT14        0      0      DISABLED  *
EXT3         0      0      DISABLED  *
EXT4         0      0      DISABLED  *
* = STP turned off for this port.

------------------------------------------------------------------
Spanning Tree Group 2: On (STP/PVST)
VLANs:  5

Current Root:            Path-Cost  Port Hello MaxAge FwdDel Aging
 6005 00:03:fd:6a:c5:80        4    EXT1    2     20     15    300

Parameters:  Priority  Hello  MaxAge  FwdDel  Aging
              65535      2      20      15     300

Port  Priority   Cost      State       Designated Bridge     Des Port
----  --------   ----    ----------  ----------------------  --------
EXT1       128      4!    FORWARDING   6005-00:03:fd:6a:c5:80     8268
EXT2       128      4!    FORWARDING   6005-00:03:fd:6a:c5:80     8268
EXT5       128      4!     BLOCKING    7005-00:03:fd:6a:c9:80     8270
EXT6       128      4!     BLOCKING    7005-00:03:fd:6a:c9:80     8270
! = Automatic path cost.

------------------------------------------------------------------
Spanning Tree Group 3: On (STP/PVST)
VLANs:  10

Current Root:            Path-Cost  Port Hello MaxAge FwdDel Aging
 600a 00:03:fd:6a:c5:80        4    EXT1    2     20     15    300

Parameters:  Priority  Hello  MaxAge  FwdDel  Aging
              65535      2      20      15     300

Port  Priority   Cost      State       Designated Bridge     Des Port
----  --------   ----    ----------  ----------------------  --------
INT2       128      4!    FORWARDING   ffff-00:11:f9:36:b7:00     8002
INT3       128      4!    FORWARDING   ffff-00:11:f9:36:b7:00     8003
EXT1       128      4!    FORWARDING   600a-00:03:fd:6a:c5:80     8268
EXT2       128      4!    FORWARDING   600a-00:03:fd:6a:c5:80     8268
EXT5       128      4!     BLOCKING    700a-00:03:fd:6a:c9:80     8270
EXT6       128      4!     BLOCKING    700a-00:03:fd:6a:c9:80     8270
! = Automatic path cost.

------------------------------------------------------------------
Spanning Tree Group 4: On (STP/PVST)
VLANs:  20

Current Root:            Path-Cost  Port Hello MaxAge FwdDel Aging
 6014 00:03:fd:6a:c9:80        4    EXT5    2     20     15    300

Parameters:  Priority  Hello  MaxAge  FwdDel  Aging
              65535      2      20      15     300

Port  Priority   Cost      State       Designated Bridge     Des Port
----  --------   ----    ----------  ----------------------  --------
INT1       128      4!    FORWARDING   ffff-00:11:f9:36:b7:00     8001
INT2       128      4!    FORWARDING   ffff-00:11:f9:36:b7:00     8002
EXT1       128      4!     BLOCKING    7014-00:03:fd:6a:c5:80     8268
EXT2       128      4!     BLOCKING    7014-00:03:fd:6a:c5:80     8268
EXT5       128      4!    FORWARDING   6014-00:03:fd:6a:c9:80     8270
```

```
EXT6       128      4!    FORWARDING   6014-00:03:fd:6a:c9:80     8270
! = Automatic path cost.

------------------------------------------------------------------
Spanning Tree Group 5: On (STP/PVST)
VLANs:  99

Current Root:            Path-Cost  Port Hello MaxAge FwdDel Aging
 6063 00:03:fd:6a:c9:80        4    EXT5    2     20     15    300

Parameters:  Priority  Hello  MaxAge  FwdDel  Aging
              65535      2      20      15     300

Port  Priority   Cost      State       Designated Bridge     Des Port
----  --------   ----    ----------  ----------------------  --------
INT4       128      4!    FORWARDING   ffff-00:11:f9:36:b7:00     8004
EXT1       128      4!     BLOCKING    7063-00:03:fd:6a:c5:80     8268
EXT2       128      4!     BLOCKING    7063-00:03:fd:6a:c5:80     8268
EXT5       128      4!    FORWARDING   6063-00:03:fd:6a:c9:80     8270
EXT6       128      4!    FORWARDING   6063-00:03:fd:6a:c9:80     8270
! = Automatic path cost.

------------------------------------------------------------------
Spanning Tree Group 16: On (STP/PVST)
VLANs:  4095

Current Root:            Path-Cost  Port Hello MaxAge FwdDel Aging
 8000 00:11:f9:36:b7:00        0  (null)    2     20     15    300

Parameters:  Priority  Hello  MaxAge  FwdDel  Aging
              32768      2      20      15     300

Port  Priority   Cost      State       Designated Bridge     Des Port
----  --------   ----    ----------  ----------------------  --------
INT1         0      0     FORWARDING *
INT2         0      0     FORWARDING *
INT3         0      0     FORWARDING *
INT4         0      0     FORWARDING *
INT5         0      0     FORWARDING *
INT6         0      0     FORWARDING *
INT7         0      0     FORWARDING *
INT8         0      0      DISABLED  *
INT9         0      0      DISABLED  *
INT10        0      0      DISABLED  *
INT11        0      0      DISABLED  *
INT12        0      0      DISABLED  *
INT13        0      0      DISABLED  *
INT14        0      0      DISABLED  *
MGT1         0      0     FORWARDING *
MGT2         0      0      DISABLED  *
* = STP turned off for this port.
```

For Cisco IOS devices, some commands you can use to verify the desired configuration and operation of the equipment are:

- **show int status**
- **show etherchannel port**
- **show lacp internal**
- **show spanning-tree**

Design or topology remarks

Although this topology is very robust and offers high redundancy, newer topologies do exist. The drawback of this topology is the convergence time after a connection-loss or a switch-failure. Our tests showed that as long as only one link of an aggregated port group is going down, the switchover time is around 1 second, which is fairly good. In contrast, when the whole aggregation (both links) or one of the switches itself go down, it takes 30 seconds to converge, which might be unacceptable for some applications. This convergence time is the natural limitation of IEEE 802.1D Spanning Tree: 15 seconds Listening-state and 15 seconds Learning-state before a port reaches Forwarding state.

It is good that in this topology we do not have to rely on trunk failover, which will not be supported on LACP trunks until the next (1.1) software release. This topology also does not need BladeServer NIC Teaming, although the BASP Teaming function should be considered with regards to its high availability capabilities.

As seen in 7.6, “Basic Layer 2 entry topology” on page 69, a better convergence time is reached with our basic topology but only by interacting with trunk failover and active or standby NIC teaming. Other ways to reduce the time that is required to recover from a topology change include recent enhancements to the Spanning Tree Protocol, such as Rapid Spanning Tree (IEEE 801.1w) , which is discussed in the next section or, alternatively, a Layer 3-based solution, which is discussed in 7.8, “Layer 3 topology sample configurations” on page 108.

Full configuration snapshots

Example 7-13 shows the output of **`/c/dump`** for the GbESM_1, and Example 7-14 on page 92 shows the output of **`show running-conf`** from the Core1 Cisco3560 Switch. See the differences to the configurations of GbESM_2 and Core2 as comments in the relevant lines respectively.

The configurations are valid for copper and fiber infrastructure, it makes no difference to the GbESM. For the uplink switches the difference is, that for fiber connections the GBIC ports G0/25 and G0/26 have been used as IEEE 802.1Q tagged ports (not aggregated) and G0/27 as interconnection of the Cisco Switches.

> **Important:** When looking at the configuration commands presented, commands toward the left margin are for Core1 or GbESM_1 switches. Where the GbESM_2 or Core2 switch configurations differ, the commands presented within comment syntax (**/*** for GbESM, **!** for Cisco, # for Extreme) are for the counterpart switch of the same type, GbESM_2 or Core2.

Example 7-13 Advanced Layer 2 Topology with PVST configuration for the GbESMs

```
/* Version 1.0.1.6,  Base MAC address 00:11:f9:36:b7:00
/* GbESM_1
/c/sys
      hprompt ena
/c/sys/access/user/uid 1
      name "USERID"
      pswd "cbd31b7a4b020a2a86b6f2b388b2a9d8e8b4271b97d91dc22045f70228a24127"
      ena
      cos admin
/c/sys/ssnmp
      name "GbESM_1”
/*     name "GbESM_2" on GbESM_2
/c/port INT1
      pvid 20
```

```
/c/port INT3
      pvid 10
/*     pvid 20 on GbESM_2
/c/port INT4
      pvid 99
/c/port EXT1
      tag ena
      pvid 5
/c/port EXT2
      tag ena
      pvid 5
/c/port EXT5
      tag ena
      pvid 5
/c/port EXT6
      tag ena
      pvid 5
/c/l2/vlan 1
   def INT1 INT2 INT3 INT4 INT5 INT6 INT7 INT8 INT9 INT10 INT11 INT12 INT13 INT14 EXT3 EXT4
/c/l2/vlan 5
      ena
      name "Native"
      def EXT1 EXT2 EXT5 EXT6
/c/l2/vlan 10
      ena
      name "VLAN_Green"
      def INT2 INT3 EXT1 EXT2 EXT5 EXT6
/*     def INT2 EXT1 EXT2 EXT5 EXT6 on GbESM_2
/c/l2/vlan 20
      ena
      name "VLAN_Red"
      def INT1 INT2 EXT1 EXT2 EXT5 EXT6
/*     def INT1 INT2 INT3 EXT1 EXT2 EXT5 EXT6 on GbESM_2
/c/l2/vlan 99
      ena
      name "MGMT"
      def INT4 EXT1 EXT2 EXT5 EXT6
/c/l2/stg 1/off
/c/l2/stg 2/clear
/c/l2/stg 2/add 5
/c/l2/stg 3/clear
/c/l2/stg 3/add 10
/c/l2/stg 4/clear
/c/l2/stg 4/add 20
/c/l2/stg 5/clear
/c/l2/stg 5/add 99
/c/l2/lacp/port EXT1
      mode active
/c/l2/lacp/port EXT2
      mode active
      adminkey 17
/c/l2/lacp/port EXT5
      mode active
/c/l2/lacp/port EXT6
      mode active
      adminkey 21
/c/l3/if 99
      ena
      addr 10.99.0.243
/*     addr 10.99.0.244 on GbESM_2
```

```
      mask 255.255.255.0
      broad 10.99.0.255
      vlan 99
/c/l3/gw 1
      ena
      addr 10.99.0.245
/c/l3/gw 2
      ena
      addr 10.99.0.246
/c/l3/frwd/off
      dirbr disabled
/
script end  /**** DO NOT EDIT THIS LINE!
```

Example 7-14   Advanced Layer 2 Topology with PVST configuration for the Cisco3560 Switches

```
!
version 12.2
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname Core1
!  hostname Core2 on Core2
!
enable password cisco
!
no aaa new-model
ip subnet-zero
ip routing
no ip domain-lookup
!
!
no file verify auto
!
spanning-tree mode pvst
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
spanning-tree vlan 1-10 priority 24576
!  spanning-tree vlan 1-10 priority 28672 on Core2
spanning-tree vlan 11-4094 priority 28672
!  spanning-tree vlan 11-4094 priority 24576 on Core2
!
vlan internal allocation policy ascending
!
!
interface Port-channel1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
interface Port-channel2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
interface GigabitEthernet0/1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 1 mode active
!
interface GigabitEthernet0/2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 1 mode active
!
interface GigabitEthernet0/3
!
interface GigabitEthernet0/4
!
interface GigabitEthernet0/5
!
interface GigabitEthernet0/6
!
interface GigabitEthernet0/7
!
interface GigabitEthernet0/8
!
interface GigabitEthernet0/9
!
interface GigabitEthernet0/10
!
interface GigabitEthernet0/11
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 2 mode active
!
interface GigabitEthernet0/12
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 2 mode active
!
interface GigabitEthernet0/13
!
interface GigabitEthernet0/14
!
interface GigabitEthernet0/15
!
interface GigabitEthernet0/16
!
interface GigabitEthernet0/17
```

```
!
interface GigabitEthernet0/18
!
interface GigabitEthernet0/19
!
interface GigabitEthernet0/20
!
interface GigabitEthernet0/21
!
interface GigabitEthernet0/22
!
interface GigabitEthernet0/23
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 2,10,20,99
 switchport mode trunk
!
interface GigabitEthernet0/24
 no switchport
 ip address 9.42.171.245 255.255.255.0
!  ip address 9.42.171.246 255.255.255.0 on Core2
!
interface GigabitEthernet0/25
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 2,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
interface GigabitEthernet0/26
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 2,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
interface GigabitEthernet0/27
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 2,10,20,99
 switchport mode trunk
!
interface GigabitEthernet0/28
!
interface Vlan1
 no ip address
!
interface Vlan10
 ip address 10.10.0.245 255.255.255.0
!   ip address 10.10.0.246 255.255.255.0 on Core2
!
interface Vlan20
 ip address 10.20.0.245 255.255.255.0
!   ip address 10.20.0.246 255.255.255.0 on Core2
!
interface Vlan99
 ip address 10.99.0.245 255.255.255.0
!   ip address 10.99.0.245 255.255.255.0 on Core2
!
ip default-gateway 9.42.171.3
ip classless
ip route 0.0.0.0 0.0.0.0 9.42.171.3
ip http server
ip http secure-server
!
control-plane
!
line con 0
line vty 0 4
 password cisco
 no login
line vty 5 15
 no login
!
end
```

7.7.3 Rapid Spanning Tree IEEE 802.1w

The GbESM supports the IEEE 802.1w RSTP (Rapid Spanning Tree) standard. When this protocol is used there can be only one Spanning Tree Group. When IEEE 802.1s Multiple Spanning Tree is used, multiple spanning tree instances can be configured. The consequence of this is that VLAN 4095 (the Management VLAN) will be moved to Spanning Tree Group 1, while Spanning Tree Group 16 (the STG for the Management VLAN) is turned off if RSTP is enabled on the GbESM.

When GbESM modules were connected to switches that run Cisco Rapid-PVST, the untagged (native) VLAN looped during our tests for this sample configuration. Because the chosen topology (see Figure 7-5 on page 81) allows STP to be disabled on the BladeCenter Switch Modules, Rapid-PVST was configured on the Cisco devices, and STP was switched off completely (STG and MRST) on both GbESMs.

Prepare the Nortel GbESM_1 and GbESM_2 for Rapid Spanning Tree

Due to incompatibilities in the RSTP behavior of the GbESM and the Cisco 3560, Spanning Tree is disabled on the GbESMs and Rapid-PVST is enabled on the Cisco Switches. Remember to disconnect the devices before starting configuration changes as described in 7.5, “Base configurations for examples in this document” on page 66.

Ensure that Rapid STP is disabled:

```
/c/l2/mrst/off
```

Move all VLANs (except management VLAN 4095) to Spanning Tree Group 1 and switch it off:

```
/c/l2/stg 1/add 1 5 10 20 99
/c/l2/stg 1/off
```

Configure the Cisco Core1 and Core2 for Rapid-PVST

Change the Spanning Tree mode from the default mode to Rapid-PVST:

```
conf t
spanning-tree mode rapid-pvst
```

In our example, we divide the Root Bridge over the VLANs to share the load, Core1 should be the Root Bridge for VLANs 5 and 10:

```
spanning-tree vlan 1-10 root primary
spanning-tree vlan 11-4094 root secondary
```

Vice versa for Core2 to become the Root Bridge for VLANs 20 and 99:

```
spanning-tree vlan 1-10 root secondary
spanning-tree vlan 11-4094 root primary
```

Finally, the port costs for the connections to the GbESMs are set to a higher value to make them relatively undesirable Spanning Tree Paths. This is to ensure that the direct connection between the Core switches is preferred:

```
interface Port-channel1
 spanning-tree cost 6
interface Port-channel2
 spanning-tree cost 6
end
```

Reenable the ports

After verifying the correct cabling between the all the devices, the ports can be enabled:

- Enabling the ports on GbESM_1 and GbESM_2

```
/oper/port EXT1/ena
/oper/port EXT2/ena
/oper/port EXT5/ena
/oper/port EXT6/ena
```

- Enabling the ports on Core1 and Core2

```
conf t
int range g0/1-2, g0/11-12
no shut
```

Verify the Spanning Tree status

Verify the port status and vlan assignment by typing the **info/link** and **info/port** command on the GbESM. The commands **info/l2/stg** and **info/l2/trunk** shows the Spanning Tree status. Verify that no ports of the GbESMs are in BLOCKING status since the Spanning Tree is disabled.

For Cisco IOS devices some commands you can use to verify the desired configuration and operation of the equipment are:

- **show int status**
- **show etherchannel port**
- **show lacp internal**
- **show spanning-tree**

Check that Core1 is root for VLAN 5 and 10, while Core2 is root for VLAN 20 and 99. Thus the ports in BLOCKING status on Core1 and Core2 should differ depending on the VLAN.

Design remarks

This topology offers high redundancy with very fast failure recovery. Independent of features like trunk failover or the NIC teaming, we reached in our tests recovery times of less than a second in optimum.

A drawback of this topology is that in some cases half of the bandwidth available between the IBM @server BladeCenter and the core is not used, due to constraints of the Spanning Tree Algorithm blocking the traffic on the backup connections.

Full configuration snapshots

Example 7-15 shows the output of **/c/dump** for the GbESM_1 while Example 7-16 on page 98 shows the output of **show running-conf** from the Core1 Cisco3560 Switch. See the differences to the configurations of GbESM_2 and Core2 as comments in the relevant lines respectively.

The configurations are valid for copper and fiber infrastructure, it makes no difference to the GbESM. For the uplink switches the difference is, that for fiber connections the GBIC ports G0/25 and G0/26 have been used as IEEE 802.1Q tagged ports (not aggregated) and G0/27 as interconnection of the Cisco Switches.

> **Important:** When looking at the configuration commands presented, commands toward the left margin are for Core1 or GbESM_1 switches. Where the GbESM_2 or Core2 switch configurations differ, the commands presented within comment syntax (**/*** for GbESM, **!** for Cisco, # for Extreme) are for the counterpart switch of the same type, GbESM_2 or Core2.

Example 7-15 Advanced Layer 2 Topology with STP turned off for the GbESMs

```
/* Version 1.0.1.6,  Base MAC address 00:11:f9:36:b7:00
/* GbESM_1
/c/sys
      hprompt ena
/c/sys/access/user/uid 1
      name "USERID"
      pswd "5754473e1340022a1e37a6f7d0f0a1d8aa8b437633b51ae68eb6d669ce9fd9f7"
      ena
      cos admin
/c/sys/ssnmp
      name "GbESM_1"
/*     name "GbESM_2" on GbESM_2
/c/port INT1
      pvid 20
/c/port INT3
      pvid 10
/*     pvid 20 on GbESM_2
/c/port INT4
      pvid 99
/c/port EXT1
      tag ena
      pvid 5
/c/port EXT2
      tag ena
      pvid 5
/c/port EXT5
      tag ena
      pvid 5
/c/port EXT6
      tag ena
      pvid 5
/c/l2/vlan 1
   def INT1 INT2 INT3 INT4 INT5 INT6 INT7 INT8 INT9 INT10 INT11 INT12 INT13 INT14 EXT3 EXT4
/c/l2/vlan 5
      ena
      name "Native"
      def EXT1 EXT2 EXT5 EXT6
/c/l2/vlan 10
      ena
      name "VLAN_Green"
      def INT2 INT3 EXT1 EXT2 EXT5 EXT6
/*     def INT2 EXT1 EXT2 EXT5 EXT6 on GbESM_2
/c/l2/vlan 20
      ena
      name "VLAN_Red"
      def INT1 INT2 EXT1 EXT2 EXT5 EXT6
/*     def INT1 INT2 INT3 EXT1 EXT2 EXT5 EXT6 on GbESM_2
/c/l2/vlan 99
      ena
      name "MGMT"
      def INT4 EXT1 EXT2 EXT5 EXT6
/c/l2/stg 1/off
/c/l2/stg 1/clear
/c/l2/stg 1/add 1 5 10 20 99
/c/l2/lacp/port EXT1
      mode active
/c/l2/lacp/port EXT2
      mode active
      adminkey 17
/c/l2/lacp/port EXT5
      mode active
/c/l2/lacp/port EXT6
      mode active
      adminkey 21
/c/l3/if 99
      ena
      addr 10.99.0.243
/*     addr 10.99.0.244 on GbESM_2
      mask 255.255.255.0
      broad 10.99.0.255
      vlan 99
/c/l3/gw 1
      ena
      addr 10.99.0.245
/c/l3/gw 2
      ena
      addr 10.99.0.246
/c/l3/frwd/off
      dirbr disabled
/
script end  /**** DO NOT EDIT THIS LINE!
```

Example 7-16   Advanced Layer 2 Topology with Rapid-PVST configuration for the Cisco3560 Switches

```
!
version 12.2
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname Core1
!  hostname Core2 on Core2
!
enable password cisco
!
no aaa new-model
```

```
ip subnet-zero
ip routing
no ip domain-lookup
!
!
no file verify auto
!
spanning-tree mode rapid-pvst
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
spanning-tree vlan 1-10 priority 24576
!  spanning-tree vlan 1-10 priority 28672 on Core2
spanning-tree vlan 11-4094 priority 28672
!  spanning-tree vlan 11-4094 priority 24576 on Core2
!
vlan internal allocation policy ascending
!
!
interface Port-channel1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 spanning-tree cost 100
!
interface Port-channel2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 spanning-tree cost 100
!
interface GigabitEthernet0/1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 1 mode active
!
interface GigabitEthernet0/2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 1 mode active
!
interface GigabitEthernet0/3
!
interface GigabitEthernet0/4
!
interface GigabitEthernet0/5
!
interface GigabitEthernet0/6
!
interface GigabitEthernet0/7
!
```

```
interface GigabitEthernet0/8
!
interface GigabitEthernet0/9
!
interface GigabitEthernet0/10
!
interface GigabitEthernet0/11
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 2 mode active
!
interface GigabitEthernet0/12
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
 channel-group 2 mode active
!
interface GigabitEthernet0/13
!
interface GigabitEthernet0/14
!
interface GigabitEthernet0/15
!
interface GigabitEthernet0/16
!
interface GigabitEthernet0/17
!
interface GigabitEthernet0/18
!
interface GigabitEthernet0/19
!
interface GigabitEthernet0/20
!
interface GigabitEthernet0/21
!
interface GigabitEthernet0/22
!
interface GigabitEthernet0/23
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
!
interface GigabitEthernet0/24
 no switchport
 ip address 9.42.171.245 255.255.255.0
!  ip address 9.42.171.246 255.255.255.0 on Core2
!
interface GigabitEthernet0/25
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
```

```
interface GigabitEthernet0/26
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
 switchport nonegotiate
!
interface GigabitEthernet0/27
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport trunk allowed vlan 5,10,20,99
 switchport mode trunk
!
interface GigabitEthernet0/28
!
interface Vlan1
 no ip address

interface Vlan10
 ip address 10.10.0.245 255.255.255.0
!   ip address 10.10.0.246 255.255.255.0 on Core2
!
interface Vlan20
 ip address 10.20.0.245 255.255.255.0
!   ip address 10.20.0.246 255.255.255.0 on Core2
!
interface Vlan99
 ip address 10.99.0.245 255.255.255.0
!   ip address 10.99.0.245 255.255.255.0 on Core2
!
ip default-gateway 9.42.171.3
ip classless
ip route 0.0.0.0 0.0.0.0 9.42.171.3
ip http server
ip http secure-server
!
control-plane
!
line con 0
line vty 0 4
 password cisco
 no login
line vty 5 15
 no login
!
end
```

7.7.4 Multi-Spanning Tree IEEE 802.1s

During the writing of this book, it was not possible to create and test a stable configuration with the GbESM and the Cisco 3560 devices both configured to run MSTP. After our testing was completed, we learned that this is remedied in IOS version 12.2(25)SEC on the Cisco 3750. We assume that similarly numbered versions on other platforms will also include this fix.

Configurations for use of MST

The MST (802.1s) protocol, like the prestandard multiple spanning tree group capability of the GbESM, allows for multiple instances of spanning tree to run at the same time. Each spanning tree instance is associated with one or more VLANs, which must be explicitly identified as part of the configuration. It is good practice to associate VLANs with the same topology to a single instance of spanning tree. This reduces processor overhead on the GbESM and other switches in the network and it reduces the traffic generated by the BPDUs that the switches send to each other.

In general, MSTP configuration mirrors the configuration of multiple instances of classic 802.1D spanning tree which is described in 7.7.2, “Common Spanning Tree configuration - IEEE 802.1D and PVST” on page 84, except as follows:

- When multiple 802.1D instances are used to interoperate with Cisco PVST+, there is typically only one VLAN associated with each Spanning Tree Group (STG). When MSTP is used, both the GbESM and any upstream switches from any vendor must explicitly identify the VLANs associated with each MSTP instance.
- MSTP requires the configuration of a region name and revision number which *must* match for all switches if they are to recognize each other as part of the same MSTP region. In almost all cases, GbESMs should be in the same MSTP region as their upstream neighbors.
- MSTP can recognize and interoperate with switches running 802.1D (classic spanning tree) or 802.1w (rapid spanning tree). The commands under /c/l2/mrst/cist define the parameters for this interoperability. These commands also govern the interactions of boundary switches which are neighbors but are part of different MSTP regions.

The MST functionality is enabled on the GbESM by the following commands:

/c/l2/mrst/on
Enables 802.1w or 802.1s (rapid and multiple spanning tree) functionality

/c/l2/mrst/mode mstp
Selects multiple spanning tree (802.1s) instead of single rapid spanning tree (802.1w)

/c/l2/mrst/cist
Menu of commands to configure the common spanning tree defined by 802.1s, which is equivalent to spanning tree group 0 (zero).

/c/l2/stg *<group number>*
Menu of commands to configure additional instances of MSTP. This includes commands to specify bridge parameters, port parameters, and to identify the VLANs associated with this spanning tree group.

/c/l2/stg *<group number/on|off*
Enables or disables a spanning tree instance.

Key configuration commands to verify MSTP

On the GbESM, the **/info/l2/stg** command shows the status of the MSTP instances when MSTP is in use. It presents output as shown below. Note that the `common internal spanning tree` is instance 0 (zero) and is used to communicate outside of the MSTP region. Its details are obtained by the **/info/l2/cist** command.

Example 7-17 MSTP Information Display - GbESM

```
>> Main# /info/l2/stg
------------------------------------------------------------------
Spanning Tree Group 1: On (MSTP)
VLANs: 10 30

Current Root:            Path-Cost     Port     Aging
1000 00:12:7f:ea:76:00    20000         EXT1     300

Parameters: Priority Aging
```

```
32768 300

Port Prio Cost State Role Designated Bridge     Des Port
---- ---- --------- ----- ---- --------------------- --------
EXT1 128 20000 FWD   ROOT 1000-00:12:7f:ea:76:00 8003
EXT2 128 20000 DSB
-------------------------------------------------------------------
Spanning Tree Group 2: On (MSTP)
VLANs: 20 40

Current Root:             Path-Cost     Port     Aging
2000 00:12:7f:ea:76:00     20000         EXT1     300

Parameters: Priority Aging
32768 300

Port Prio Cost State Role Designated Bridge Des Port
----- ---- --------- ----- ---- --------------------- --------
EXT1 128 20000 FWD ROOT 2000-00:12:7f:ea:76:00 8003
EXT2 128 20000 DSB

>> Layer 2# /info/l2/cist
-------------------------------------------------------------------
Common Internal Spanning Tree:

VLANs: 1-9 11-19 21-29 31-39 41-4094

Current Root:             Path-Cost Port MaxAge FwdDel
3000 00:12:7f:ea:76:00 0                17     20             15

Cist Regional Root:     Path-Cost
3000 00:12:7f:ea:76:00     20000

Parameters: Priority MaxAge FwdDel Hops
32768 20 15 20
Port Prio Cost State Role Designated Bridge Des Port Hello Type
----- ---- --------- ----- ---- --------------------- -------- ----- ----
INT1 0 0 DSB *
INT2 0 0 DSB *
INT3 0 0 DSB *
INT4 0 0 DSB *
INT5 0 0 DSB *
INT6 0 0 DSB *
INT7 0 0 DSB *
INT8 0 0 DSB *
INT9 0 0 DSB *
INT10 0 0 DSB *
INT11 0 0 DSB *
INT12 0 0 DSB *
INT13 0 0 DSB *
INT14 0 0 DSB *
MGT1 0 0 FWD *
MGT2 0 0 DSB *
EXT1 128 20000 FWD ROOT 3000-00:12:7f:ea:76:00 8003 2 P2P
EXT2 128 20000 DSB
EXT3 128 20000 DSB
EXT4 128 20000 DSB
EXT5 128 20000 DSB
EXT6 128 20000 DSB
```

On the Cisco 3750, the **sh spanning-tree mst <instance>** command provides similar data as shown below. It is possible to tell that the connected ports on the GbESM and 3750 are successfully interoperating using MST because neither of the ports is displayed as being a *boundary port*. Boundary ports on switches running MSTP are ports that connect to switches running other, earlier versions of spanning tree such as RSTP or original 802.1D spanning tree. Connected ports which are parts of different MSTP regions also are boundaries, and this was observed in tests with earlier versions of Cisco firmware which did not fully support 802.1s.

Example 7-18 MSTP Information Display - Cisco

```
Switch#show spanning-tree mst 0

##### MST0 vlans mapped: 1-9,11-19,21-29,31-39,41-4094
Bridge address 0012.7fea.7600 priority 12288 (12288 sysid 0)
Root this switch for the CIST
Operational hello time 2 , forward delay 15, max age 20, txholdcount 6
Configured hello time 2 , forward delay 15, max age 20, max hops 20

Interface     Role     Sts     Cost     Prio.Nbr     Type
------------ ---- --- ------- -------- --------------------------------
Fa1/0/1        Desg     FWD    200000   128.3         P2p

##### MST1 vlans mapped: 10,30
Bridge address 0012.7fea.7600 priority 4097 (4096 sysid 1)
Root this switch for MST1

Interface     Role     Sts     Cost     Prio.Nbr     Type
---------------- ---- --- --------- -------- --------------------------------
Fa1/0/1        Desg     FWD   200000    128.3         P2p

##### MST2 vlans mapped: 20,40
Bridge address 0012.7fea.7600 priority 8194 (8192 sysid 2)
Root this switch for MST2

Interface     Role     Sts    Cost     Prio.Nbr      Type
---------------- ---- --- --------- -------- --------------------------------
Fa1/0/1     Desg       FWD     200000     128.3      P2p
```

Note that the configurations Example 7-19 were done by Nortel teams in Santa Clara and not as part of the work of the team that performed the other tests which are documented in this Redpaper. Thus, the following configurations are not based on the configurations shown earlier. The key commands for MRST (802.1s) are shown in boldface.

This configuration creates two MRST instances, STG 1 for VLANS 10 and 30 and the other for VLANS 20 and 40. STG 16, which is used exclusively for the Management Module VLAN (4095) is also active by default. The STP region name is `INTEROP`.

Example 7-19 Sample MSTP Configuration - GbESM

```
script start "Layer 2-3 Gigabit Ethernet Switch Module for IBM eServer
BladeCenter" 4 /**** DO NOT EDIT THIS LINE!
/* Configuration dump taken 0:05:23 Thu Jan 1, 2070
/* Version 1.0.1.6, Base MAC address 00:11:f9:38:1b:00
/c/port EXT1
tag ena
/c/port EXT2
tag ena
/c/l2/vlan 10
ena
```

```
name "VLAN 10"
def EXT1 EXT2
/c/l2/vlan 20
ena
name "VLAN 20"
def EXT1 EXT2
/c/l2/vlan 30
ena
name "VLAN 30"
def EXT1 EXT2
/c/l2/vlan 40
ena
name "VLAN 40"
def EXT1 EXT2
/c/l2/mrst/on
/c/l2/mrst/mode mstp
/c/l2/mrst/name INTEROP
/c/l2/stg 1/clear
/c/l2/stg 1/add 10 30
/c/l2/stg 2/clear
/c/l2/stg 2/add 20 40
/c/l2/stg 16/clear

/c/l3/if 1
ena
addr 172.16.1.3
mask 255.255.255.0
broad 172.16.1.255
/c/l3/if 10
ena
addr 172.16.10.3
mask 255.255.255.0
broad 172.16.10.255
vlan 10
/c/l3/if 20
ena
addr 172.16.20.3
mask 255.255.255.0
broad 172.16.20.255
vlan 20
/c/l3/if 30
ena
addr 172.16.30.3
mask 255.255.255.0
broad 172.16.30.255
vlan 30
/c/l3/if 40
ena
addr 172.16.40.3
mask 255.255.255.0
broad 172.16.40.255
vlan 40
/c/l3/frwd/off
dirbr disabled
/
script end /**** DO NOT EDIT THIS LINE!
```

Example 7-20 was tested by being interconnected with Example 7-18 on page 104. The same parameters for MSTP and VLANs are used. A Cisco 3750 switch is used to connect to the Nortel GbESM in the BladeCenter chassis.

Example 7-20 Sample MSTP Configuration - Cisco

```
version 12.2
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname Switch
!
enable password c
!
no aaa new-model
switch 1 provision ws-c3750-24ts
vtp domain BC
vtp mode transparent
ip subnet-zero
!
no file verify auto
!
spanning-tree mode mst
spanning-tree extend system-id
!
spanning-tree mst configuration
name INTEROP
revision 1
instance 1 vlan 10, 30
instance 2 vlan 20, 40
!
spanning-tree mst 0 priority 12288
spanning-tree mst 1 priority 4096
spanning-tree mst 2 priority 8192
!
vlan internal allocation policy ascending
!
vlan 10,20,30,40
!
!
interface FastEthernet1/0/1
switchport trunk encapsulation dot1q
switchport mode trunk
!
interface FastEthernet1/0/2
shutdown
!
interface FastEthernet1/0/3
shutdown
!
interface FastEthernet1/0/4
shutdown
!
interface FastEthernet1/0/5
shutdown
!
interface FastEthernet1/0/6
shutdown
!
```

```
interface FastEthernet1/0/7
shutdown
!
interface FastEthernet1/0/8
shutdown
!
interface FastEthernet1/0/9
shutdown
!
interface FastEthernet1/0/10
shutdown
!
interface FastEthernet1/0/11
shutdown
!
interface FastEthernet1/0/12
shutdown
!
interface FastEthernet1/0/13
shutdown
!
interface FastEthernet1/0/14
shutdown
!
interface FastEthernet1/0/15
shutdown
!
interface FastEthernet1/0/16
shutdown
!
interface FastEthernet1/0/17
shutdown
!
interface FastEthernet1/0/18
shutdown
!
interface FastEthernet1/0/19
shutdown
!
interface FastEthernet1/0/20
shutdown
!
interface FastEthernet1/0/21
shutdown
!
interface FastEthernet1/0/22
shutdown
!
interface FastEthernet1/0/23
shutdown
!
interface FastEthernet1/0/24
shutdown
!
interface GigabitEthernet1/0/1
!
interface GigabitEthernet1/0/2
!
interface Vlan1
ip address 172.16.1.10 255.255.255.0
```

```
!
interface Vlan10
ip address 172.16.10.10 255.255.255.0
!
ip classless
ip http server
ip http secure-server
!
control-plane
!
line con 0
line vty 0 4
password c
login
line vty 5 15
no login
!
end
```

7.8 Layer 3 topology sample configurations

The configuration samples in this section use Layer 3 switching (routing) in addition to Layer 2 switching (bridging, forwarding). They are based on the simple configuration that was introduced in 7.6, “Basic Layer 2 entry topology” on page 69.

Implications of using Layer 3

These Layer 3 configurations are different from the Layer 2 configurations in the following ways:

- 802.1q trunking is not required on the connections to the Core switches. Traffic from the VLANs used for the server blades (VLAN 10, VLAN 20) is sent on the uplink trunks as routed traffic and identified by its source and destination IP addresses. There is no need to identify its VLAN association.
- Additional VLANs and associated subnets are required for the connections to the Core switches. The uplink connections do not share the VLANs (10, 20) or subnets (10.10.x.x, 10.20.x.x) used for the blade servers. In some of the examples that follow only VLANs 35 and 46 are used. The VLANs and subnets used are:
  - VLAN 35: 10.35.x.x, used to connect GbESM 1 to upstream routers
  - VLAN 36: 10.36.x.x, used to connect GbESM 1 to upstream router Core2
  - VLAN 45: 10.45.x.x, used to connect GbESM 2 to upstream router Core1
  - VLAN 46: 10.46.x.x, used to connect GbESM 2 to upstream routers
- Upstream routers must be able to find the route to the blade servers. This requires either the use of explicitly configured (static) routes or the use of a dynamic routing protocol. Dynamic routing protocols enable routers to share route information with their neighbors. Examples of the use of static and dynamic routing are included in this section.
- High Availability designs using Layer 3 require different techniques than those using solely Layer 2. Examples of HA designs are included in this section.

Reasons for choosing a Layer 3 design

The reasons for choosing a Layer 3 design for blade server switching are outlined below. It is worth noting that in most discussions of network architecture, the blade servers would be connected to a so-called *server access* switch. This is a switch at the edge of the network which connects directly to servers as opposed to client (desktop, mobile computer) computers. Most discussions of network architecture recommend that server access switches be Layer 3 devices.

- Layer 3 switching keeps more traffic within the BladeCenter chassis.

  In any design where the blade servers are assigned to multiple VLANs, routing is required for servers which are not on the same VLAN to communicate with each other. (An example of such a design would be using blade servers as WebSphere® Web and application servers and placing the Web servers on one VLAN and the application servers on a different VLAN.)

  With a Layer 3 configuration on the GbESM, servers on different VLANs can communicate via the switch module which is inside the BladeCenter chassis. If only Layer 2 switching is used within the chassis, then traffic between, for example, WebSphere Web and application servers would leave the chassis on the external links and flow through one or more external devices until it reached a Layer 3 switch (router). It would then flow through one or more additional devices until it returned to the IBM @server BladeCenter, crossing the external links for a second time.

  The benefits of keeping traffic within the chassis include greater security - because there are no patch panels which can be tampered with - and lower latency, since the traffic traverses a smaller number of switches between source and destination.

- Layer 3 switching allows more efficient use of external connections

  The key issue here is that use of Layer 3 allows the network to run without the use of Spanning Tree Protocol (STP). Spanning tree works by blocking links which would create a topology which includes a loop; connections from a GbESM to two or more upstream switches which are connected to each other fall into this category. The consequence is that up to half of the links from the GbESM would be *blocked* during normal operations and would not carry traffic. Only if the active link(s) failed would the blocked links be used to carry traffic.

  Layer 3 routing not only allows all of the uplinks to be active but also allows the GbESM to send traffic to a given destination on the best path to that destination.

7.8.1 Layer 3 sample configuration with static routing and VRRP

This example is an extension to the basic configuration described in 7.6, “Basic Layer 2 entry topology” on page 69, providing a mesh topology between the two GbESM switches and the upstream Core switches. Trunking (link aggregation) is used in this configuration in the same way it was used in the advanced Layer 2 configurations shown in 7.7, “Advanced Layer 2 topology sample configurations” on page 80.

See Figure 7-6 on page 111 for a diagram of the topology used in this example.

VLAN 35 is configured and connects both of the GbESM switches to Core switch 1 (address .245). VLAN 46 is configured to connect both of the GbESMs to Core switch 2 (address .246). This configuration uses VRRP and hot standby to provide high availability.

VRRP is configured on VLANs 10 and 20 which face towards the server blades and VLANs 35 and 46 which face the Core routers. VRRP uses an IP address which is shared between the two GbESM modules. At any moment, one of the GbESM modules is the VRRP master, and only the master responds to the shared address. In this sample, GbESM 1 is configured with a higher priority than GbESM 2 and it will therefore be the master if it is operational.

Hot-standby is an option of VRRP which works similarly to the trunk failover feature used with the Layer 2 configurations. When VRRP is configured with hot-standby, the standby switch will disable all the internal ports within the chassis, triggering the NIC teaming driver to use the ports which connect to the other switch.

Note: It is possible to use VRRP without the hot-standby feature. Such a design survives the failure of the primary switch module in the same way as it would with hot-standby enabled. In some circumstances the failure of the uplink ports (or upstream switch) is not protected against without the use of hot-standby. It is always advisable to test a proposed configuration to verify that it will protect against all the failure modes that you intend it to.

Merits of this configuration

This configuration provides strong High Availability features. Its behavior in response to common failure modes is as follows:

- Switch failure, where traffic flows via the other switch; NIC teaming and VRRP ensure routing to the appropriate places. Servers default gateway is still valid and static routes pointing to the externally facing VRRP addresses are still valid.
- Uplink failure, where traffic flows either via the other switch and through a cross-over link connecting the two GbESM switches or, with hot standby and tracking, disables the internal ports and trigger NIC teaming thus sending all traffic through the second switch.

Summary of disconnect procedure to be performed for each example

When performing initial configurations or making changes to existing configurations that might have an impact on a routed network, it is recommended that you leave connections un-cabled, or shut down, prior to making the configuration changes. This will reduce the likelihood of any routing loops or other transient behaviors which might disrupt other parts of the network.

- Shut down the ports on GbESM1 and GbESM2

```
/oper/port EXT1/dis
/oper/port EXT2/dis
/oper/port EXT5/dis
/oper/port EXT6/dis
```

- Shut down the ports on Core1 and Core2

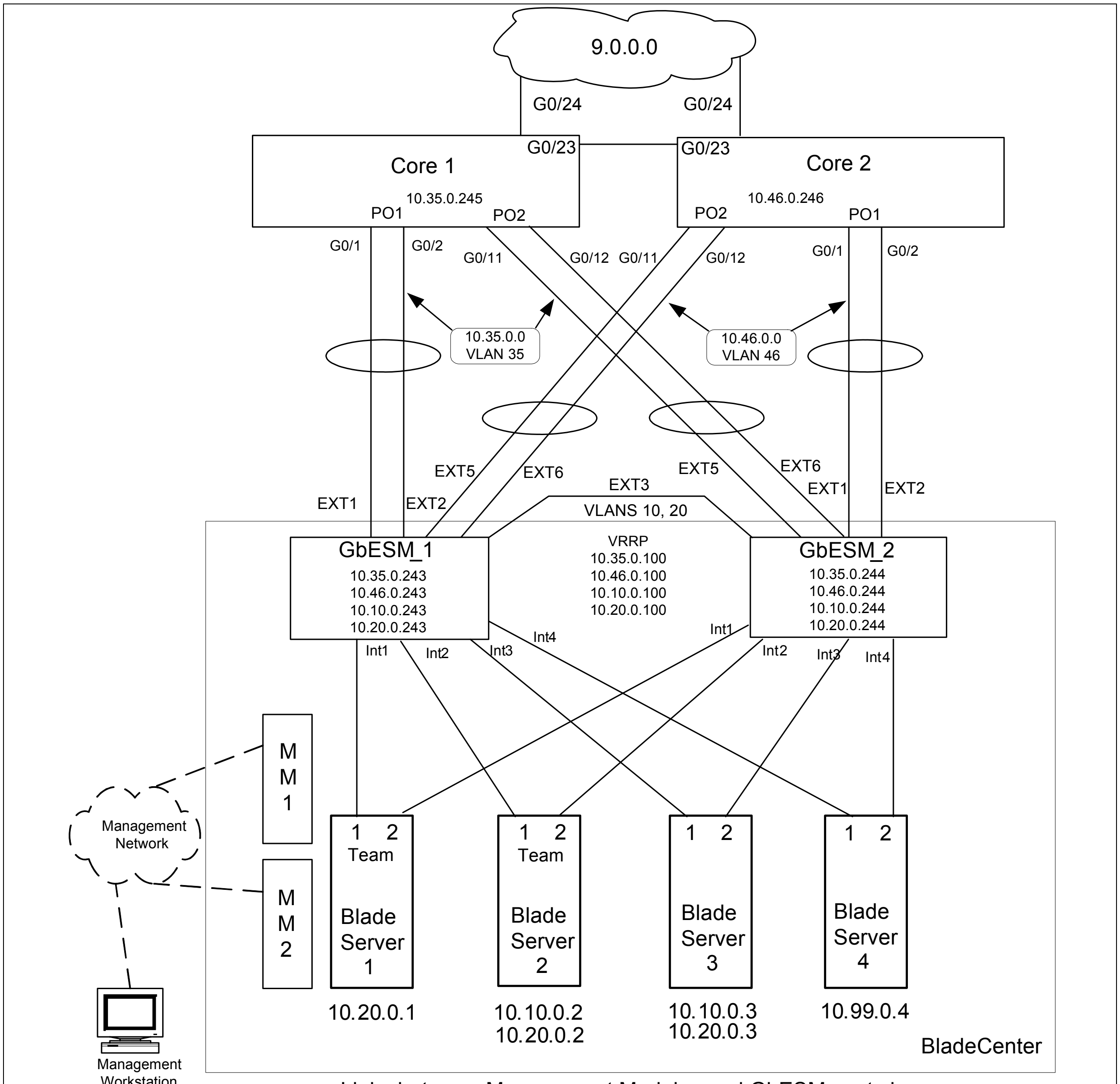
```
conf t
int range g0/1-2, g0/11-12
shut
```

Summary of IP Addressing used in this sample

Table 7-1 summarizes the IP addresses used in this sample. Figure 7-6 also illustrates these.

Table 7-1 IP Addresses for Layer 3 Sample Configuration (Static routing, VRRP)

Switch	VLAN 35	VLAN 46	VLAN 10	VLAN 20
GbESM 1	10.35.0.243	10.46.0.243	10.10.0.243	10.20.0.243
GbESM 2	10.35.0.244	10.46.0.244	10.10.0.244	10.20.0.244
Core 1	10.35.0.245	none	none	none
Core 2	none	10.46.0.246	none	none
VRRP - GbESMs	10.35.0.100	10.46.0.100	10.10.0.100	10.20.0.100



Links between Management Modules and GbESMs not shown

Figure 7-6 Layer 3 sample configuration (static routing, VRRP)

Configure VLANs and subnets for the uplink ports

The ports which connect the GbESMs to the Core switches are assigned to their own VLANs and associated subnets are configured. This process is done on the GbESM modules as well as the Core switches. Access to the GbESMs via the management module or the serial console port may be required to successfully achieve this while the external ports are disconnected.

This configuration uses the same port aggregation (trunking) as the Layer 2 configurations: Ports EXT1 and EXT2 on each GbESMs are trunked statically and connect to ports Gi0/1 and Gi0/2 on the corresponding Core switch. Ports EXT5 and EXT6 on the GbESMs use LACP and connect to ports Gi0/11 and Gi0/12 on the opposite (GbESM1 to Core2, GbESM2 to Core1) Core switch. The GbESMs also require default gateways on the VLANs which connect them to the Core switches. These gateways will be 10.35.0.245 and 10.46.0.246 and are shown in the configuration text below.

The two GbESMs and the two Core switches each use configurations (shown in Example 7-21 and Example 7-22 on page 114) which are identical except where noted.

> **Important:** When looking at the configuration commands presented, commands toward the left margin are for Core1 or GbESM_1 switches. Where the GbESM_2 or Core2 switch configurations differ, the commands presented within comment syntax (**/*** for GbESM, **!** for Cisco, # for Extreme) are for the counterpart switch of the same type, GbESM_2 or Core2.

Example 7-21 Layer 3 VLANs and ports for GbESM switches

```
/* ensure L3 functions are enabled

/cfg/l3/frwd/on

/* configure default gatways pointing at core switches
/cfg/l3/gw 1
   addr 10.35.0.245
ena
/cfg/l3/gw 2
addr 10.46.0.246
ena

/* create trunk group for cross connect to second core switch
/cfg/l2/trunk 2
   add EXT5
   add EXT6
   ena

/* configure ports and build VLANs
/cfg/l2/vlan 35/ena
/cfg/l2/vlan 46/ena

/cfg/port ext1/tag d
/cfg/port ext2/tag d
/cfg/port ext5/tag d
/cfg/port ext6/tag d

/cfg/port ext3/tag e
/* cross-over to other GbESM needed for VRRP

/* for GbESM 1
/cfg/l2/vlan 35/def ext1 ext2
```

```
/cfg/l2/vlan 46/def ext5 ext6

/* for GbESM 2
/* /cfg/l2/vlan 35/def ext5 ext6
/* /cfg/l2/vlan 46/def ext1 ext2

/* both GbESMs use the below...
/cfg/l2/vlan 10/add ext3
/cfg/l2/vlan 20/add ext3


/* create L3 interfaces on VLANs
/* these are for the subnets where the servers are
/cfg/l3/if 10
   ena
   vlan 10
   addr 10.10.0.243
/* addr 10.10.0.244 on GbESM 2
   mask 255.255.255.0
/cfg/l3/if 20
   ena
   vlan 20
   addr 10.20.0.243
/* addr 10.20.0.244 for GbESM 2
   mask 255.255.255.0

/* these are for the uplink subnets
/cfg/l3/if 35
   ena
   vlan 35
   addr 10.35.0.243 --
/* addr 10.35.0.244 for GbESM 2
   mask 255.255.255.0

/cfg/l3/if 46
   ena
   vlan 46
   addr 10.46.0.243
/* addr 10.46.0.244 for GbESM 2
   mask 255.255.255.0

   /* clean up VLAN 1 - remove unused ports
   /c/l2/vlan 1
      rem EXT1
      rem EXT2
      rem EXT3
      rem EXT5
      rem EXT6
   /* clean up VLAN 5 - unused in this configuration
   /c/l2/vlan 5
      del
```

Example 7-22 Layer 3 VLANs and ports for Core switches

```
Configuration for Core1 (address .245)
   ! first of all enable layer 3, disable STP/PVST, clean up unused
   ip routing
   no spanning-tree vlan 1-4094
   no vlan 10
   no vlan 20
   no vlan 5
   !
   ! assign ports and VLANs
   interface range g0/1-2
   channel-group 1 mode on

   interface range Po1, g0/1-2
   switchport access vlan 35
   switchport mode access

   interface range g0/11-12
   channel-group 2 mode on

   interface range Po2, g0/11-12
   switchport access vlan 35
   switchport mode access

   interface vlan 35
   ip address 10.35.0.245 255.255.255.0

Configuration for Core2 (address .246)
   ! first of all enable Layer 3, disable spanning tree, clean up
   ip routing
   no spanning-tree vlan 1-4094
   no vlan 10
   no vlan 20
   no vlan 5
   !
   interface range g0/1-2
   channel-group 1 mode on

   interface range Po1, g0/1-2
   switchport access vlan 46
   switchport mode access

   interface range g0/11-12
   channel-group 2 mode on

   interface range Po2, g0/11-12
   switchport access vlan 46
   switchport mode access

   interface vlan 46
   ip address 10.46.0.246 255.255.255.0
```

VRRP configuration

The VRRP configuration establishes the .100 addresses on VLANs 10, 20, 35, and 46. The server blades and Core routers must also be configured to be aware of these addresses.

Configuration for GbESMs

The configuration in Example 7-23 enables VRRP globally and creates four instances of VRRP, one for each of the VLANs 10, 20, 35, and 46. VRRP priority is set to 101 instead of the default of 100 to ensure that GbESM 1 will be the VRRP master for all four VLANs. The two switches use identical configurations except where noted.

> **Note:** The **`vr`** numbers configured in the sample are unique to the switch and are purely local instance numbers. The **`vrid`** numbers must be the same on all switches which work together and must be unique to that group of switches on the VLAN where they are used.

Example 7-23 VRRP Configuration for GbESMs

```
/* turn VRRP on and then create VRRP instances
/cfg/l3/vrrp/on

/cfg/l3/vrrp/vr 1
vrid 1
if 10
addr 10.10.0.100
prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2
ena

/cfg/l3/vrrp/vr 2
vrid 2
if 20
addr 10.20.0.100
prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2

ena

/cfg/l3/vrrp/vr 3
vrid 3
if 35
addr 10.35.0.100
prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2

ena

/cfg/l3/vrrp/vr 4
vrid 4
if 46
addr 10.46.0.100
prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2

ena
```

Configuration of Hot-standby for VRRP

The commands in Example 7-24 add the hot-standby feature to the VRRP configuration. Hot-standby is used in this sample with *interface tracking.*

VRRP tracking allows a device to dynamically adjust its priority based on the availability of other resources, and allow a standby switch to become master if the current master loses access to those resources. Interface tracking counts the number of active IP interfaces and increments the VRRP priority 2 units for each by default. The effect of use of interface tracking in the sample configuration is to trigger VRRP failover and NIC teaming failover if one of the uplink trunks connecting the master GbESM to the Core switches fails. Tracking with hot-standby is applied to a VRRP *group* rather than to individual instances.

> **Note:** The VRRP tracking features of the GbESM are very flexible and allow granular control of when failover is triggered. There are many other possibilities besides what is shown in this sample. Planning and testing of these features is essential.

Note that a VLAN is created solely for the connection between the two switches. This is to avoid triggering the hot-standby failover inappropriately. The same configuration is used for both GbESM modules except where noted.

Example 7-24 VRRP Hot-standby configuration

```
/* create VLAN 50 and associated interface
/c/l2/vlan 50
   ena
   def ext3

/c/port ext3/pvid 50

/c/l3/if 50
   addr 10.50.0.243
/* addr 10.50.0.244 for GbESM2
   mask 255.255.255.0
   vlan 50
   ena

/* enable hot standby and VRRP group
/c/l3/vrrp/hotstan ena

/c/l3/vrrp/group
   ena
   if 50
   prio 101
/* above only on GbESM1; defaults to 100 on GbESM 2
   track
   /* track ip interfaces on VLANs - they go down if no active ports are found
      ifs ena
```

Configuration on Core switches to work with VRRP

The Core switches in this sample are not running VRRP. It is possible to develop a similar configuration using VRRP or Cisco's proprietary HSRP to allow the Core routers to present a shared virtual address to the GbESMs. This virtual address is configured as the GbESM's default gateway.

The necessary static routes for the Core switch 1 are:

```
ip route 10.10.0.0 255.255.255.0 10.35.0.100
ip route 10.20.0.0 255.255.255.0 10.35.0.100
```

Because Core switch 2 uses VLAN 46 rather than 35 to send traffic to the GbESMs, its static routes are different as shown below:

```
ip route 10.10.0.0 255.255.255.0 10.46.0.100
ip route 10.20.0.0 255.255.255.0 10.46.0.100
```

Configuration on server blades to work with VRRP

The server blades default gateway must be set to work with the shared VRRP addresses configured on the GbESMs. This is done using by opening the properties dialog of the network object in Windows and then the properties dialog of the TCP/IP protocol:

- Blade server 1 is only on VLAN 20 and will have a default gateway of 10.20.0.100.
- Blade server 2 is on VLANs 10 and 20 using NIC teaming and tagging; each of the VLAN network objects will have its own default gateway. VLAN 10 will use 10.10.0.100; VLAN 20 will use 10.20.0.100.
- Blade server 3 has the NIC attached to GbESM 1 on VLAN 10 and the NIC attached to GbESM 2 on VLAN 20 and will use the same addresses as blade server 2.

Reconnect procedure when configuration is ready

If the ports were disabled using the commands shown above, then re-enable them using the following commands. If they were merely unplugged, they can simply be plugged back in.

- Enable the ports on GbESM1 and GbESM2

```
/oper/port EXT1/ena
/oper/port EXT2/ena
/oper/port EXT3/ena
/oper/port EXT5/ena
/oper/port EXT6/ena
```

- Enable the ports on Core1 and Core2

```
conf t
int range Po1-2
no shut
int range g0/1-2, g0/11-12
no shut
```

Verifying operation of the configuration

You should use the commands in Example 7-25 on page 118 and Example 7-26 on page 119 to verify that the configuration is functioning as intended.

To verify that ports are successfully connected to the other device and are in the prover VLAN, use the **/i/port, /i/l2/trunk,** and **/i/link** commands. The ports should be shown as being up. The trunks should be forwarding, and VLAN membership should be as shown in the configuration with tagging off.

Example 7-25 Verifying operation of the Layer 3 configuration

```
/i/port
Alias  Port  Tag  FAST  PVID       NAME               VLAN(s)
-----  ----  ---  ----  ----  -------------  -----------------------
INT1     1    y    n     20  INT1               1 4095   20
INT2     2    y    n      1  INT2               1 4095   10   20
INT3     3    y    n     10  INT3               1 4095   10
INT4     4    y    n      1  INT4               1 4095
INT5     5    y    n      1  INT5               1 4095
INT6     6    y    n      1  INT6               1 4095
INT7     7    y    n      1  INT7               1 4095
INT8     8    y    n      1  INT8               1 4095
INT9     9    y    n      1  INT9               1 4095
INT10   10    y    n      1  INT10              1 4095
INT11   11    y    n      1  INT11              1 4095
INT12   12    y    n      1  INT12              1 4095
INT13   13    y    n      1  INT13              1 4095
INT14   14    y    n      1  INT14              1 4095
MGT1    15    y    n   4095  MGT1            4095
MGT2    16    y    n   4095  MGT2            4095
EXT1    17    n    n     35  EXT1              35
EXT2    18    n    n     35  EXT2              35
EXT3    19    y    n     50  EXT3              10   20   50
EXT4    20    n    n      1  EXT4               1
EXT5    21    n    n     36  EXT5              36
EXT6    22    n    n     36  EXT6              36

>> GbESM_1 - Information# /i/link
------------------------------------------------------------------
Alias   Port   Speed    Duplex     Flow Ctrl      Link
-----   ----   -----   --------  --TX-----RX--   ------
INT1      1     1000     full     yes    yes       up
INT2      2     1000     full     yes    yes       up
INT3      3     1000     full     yes    yes       up
INT4      4     1000     full     yes    yes       up
INT5      5     1000     full     yes    yes       up
INT6      6     1000     full     yes    yes       up
INT7      7     1000     full     yes    yes       up
INT8      8     1000     full     yes    yes      down
INT9      9     1000     full     yes    yes      down
INT10    10     1000     full     yes    yes      down
INT11    11     1000     full     yes    yes      down
INT12    12     1000     full     yes    yes      down
INT13    13     1000     full     yes    yes      down
INT14    14     1000     full     yes    yes      down
MGT1     15      100     full     yes    yes       up
MGT2     16      100     full     yes    yes    disabled
EXT1     17     1000     full      no     no       up
EXT2     18     1000     full      no     no       up
EXT3     19     1000     full     yes    yes       up
EXT4     20      any      any     yes    yes      down
EXT5     21     1000     full      no     no       up
EXT6     22     1000     full      no     no       up

>> GbESM_1 - Information# /i/l2/trunk
Trunk group 1: Enabled
failover dis, port state:
 EXT1: STG  1 forwarding
 EXT2: STG  1 forwarding
```

```
Trunk group 24: Enabled
port state:
 EXT5: STG  1 forwarding
 EXT6: STG  1 forwarding
```

VRRP operation can be verified using the **/i/l3/vrrp** command. The **/i/l3/ip** command provides general information about TCP/IP. The four VRRP instances should be up and they should show as master on GbESM 1 and standby on GbESM 2. The default gateways shown in the **/i/l3/ip** command should be up.

Example 7-26 Verifying VRRP status

```
>> GbESM_1 - Information# /i/l3/vrrp
VRRP information: (group priorities 113): hotstan, master
   1: vrid   1, 10.10.0.100,     if 10, renter, prio 113, master
   2: vrid   2, 10.20.0.100,     if 20, renter, prio 113, master
   3: vrid   3, 10.35.0.100,     if 35, renter, prio 113, master
   4: vrid   4, 10.46.0.100,     if 46, renter, prio 113, master
>> GbESM_1 - Information# /i/l3/ip
IP information:
  AS number 0

Interface information:
 10: 10.10.0.243     255.255.255.0   10.10.0.255,     vlan 10, up
 20: 10.20.0.243     255.255.255.0   10.20.0.255,     vlan 20, up
 35: 10.35.0.243     255.255.255.0   10.35.0.255,     vlan 35, up
 36: 10.36.0.243     255.255.255.0   10.36.0.255,     vlan 36, up
128: 9.42.171.243    255.255.255.0   9.42.171.255,    vlan 4095, up

Default gateway information: metric strict
132: 9.42.171.242,    vlan 4095, up

Current IP forwarding settings: ON, dirbr disabled


Current network filter settings:
  none

Current route map settings:
```

To verify end-to-end connectivity, the **ping** and **traceroute (or tracert)** commands should be used. It should be possible to ping the server blades from all of the switches and from a mobile computer that is connected to the network. This should continue to work if GbESM 1 is removed or powered off, validating the proper function of VRRP.

Complete configuration snapshots

Example 7-27 and Example 7-28 on page 122 show complete configuration files for the GbESM and upstream Core switches. One configuration for each type of switch is shown with notes where the second switch of the same type has a different configuration.

Example 7-27 Layer 3 GbESM configuration - static routing and VRRP

```
>> GbESM_1 - Main# /c/d
script start "Layer 2-3 Gigabit Ethernet Switch Module for IBM eServer BladeCenter" 4
/**** DO NOT EDIT THIS LINE!
/* Configuration dump taken  0:09:11 Thu Jan  1, 2070
/* Version 1.0.1.6,  Base MAC address 00:11:f9:36:b7:00
/* GbESM_1
/c/sys
        hprompt ena
/c/sys/access/user/uid 1
        name "USERID"
        pswd "8348a3908340a280be85e2f340f00172d60dd8c46734142520ece56dd882ccee"
        ena
        cos admin
/c/sys/ssnmp
        name "GbESM_1"
/c/port INT1
        pvid 20
/c/port INT2
        pvid 10
/c/port INT3
        pvid 10
/* -- pvid 20 on GbESM 2
/c/port EXT1
        pvid 35
/c/port EXT2
        pvid 35
/c/port EXT3
        tag ena
        pvid 50
/c/port EXT5
        pvid 46
/c/port EXT6
        pvid 46
/c/l2/vlan 1
        def INT1 INT2 INT3 INT4 INT5 INT6 INT7 INT8 INT9 INT10 INT11 INT12 INT13 INT14 EXT4
/c/l2/vlan 10
        ena
        name "VLAN_Green"
        def INT2 INT3 EXT3
/c/l2/vlan 20
        ena
        name "VLAN_Red"
        def INT1 INT2 EXT3
/* def INT1 INT2 INT3 EXT3 on GbESM 2
/c/l2/vlan 35
        ena
        name "VLAN 35"
        def EXT1 EXT2
/c/l2/vlan 46
        ena
        name "VLAN 46"
        def EXT5 EXT6
```

```
/* the below is used solely to provide an unused VLAN as PVID for the crossover

/c/l2/vlan 50
      ena
      def EXT3

/c/l2/stg 1/off
/c/l2/stg 1/clear
/c/l2/stg 1/add 1 10 20 35 46 50
/c/l2/trunk 1
        ena
        failovr dis
        add EXT1
        add EXT2
/c/l2/trunk 2
      ena
      failovr dis
      add EXT5
      add EXT6
/c/l3/if 10
        ena
        addr 10.10.0.243
/* addr 10.10.0.244 on GbESM 2
        mask 255.255.255.0
        broad 10.10.0.255
        vlan 10
/c/l3/if 20
        ena
        addr 10.20.0.243
/* addr 10.20.0.244 on GbESM 2
        mask 255.255.255.0
        broad 10.20.0.255
        vlan 20
/c/l3/if 35
        ena
        addr 10.35.0.243
/* addr 10.35.0.244 on GbESM 2
       mask 255.255.255.0
        broad 10.35.0.255
        vlan 35
/c/l3/if 46
        ena
        addr 10.46.0.243
/*        addr 10.46.0.244 on GbESM 2
        mask 255.255.255.0
        broad 10.46.0.255
        vlan 46

/c/l3/if 50
      ena
      addr 10.50.0.243
/*                 244 on GbESM 2
      mask 255.255.255.0
      vlan 50

/c/l3/gw 1
        ena
        addr 10.35.0.245
/c/l3/gw 2
        ena
        addr 10.46.0.246
/c/l3/vrrp/on
/c/l3/vrrp/vr 1
        ena
        vrid 1
        if 10
        prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2
        addr 10.10.0.100
/c/l3/vrrp/vr 2
        ena
        vrid 2
        if 20
        prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2
        addr 10.20.0.100
/c/l3/vrrp/vr 3
        ena
        vrid 3
        if 35
        prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2
        addr 10.35.0.100
/c/l3/vrrp/vr 4
        ena
        vrid 4
        if 46
        prio 101
/* the above is only on GbESM 1; defaults to 100 on GbESM 2
        addr 10.46.0.100
/
/c/l3/vrrp
      hotstan ena
/c/l3/vrrp/group
      ena
      if 50
      /* on GbESM 1 only - default to 100 on GbESM 2
      prio 101
      track
         ifs ena

script end  /**** DO NOT EDIT THIS LINE!
```

Example 7-28   Layer 3 Core switch configuration - static routing

```
! note that VLANs must be created with the Vlan <x> command in “config t” mode
! these statements are not displayed by the “sh run” command

Core1#sh run
Building configuration...

Current configuration : 3019 bytes
!
version 12.2
service config
no service pad
service timestamps debug uptime
service timestamps log uptime
```

```
no service password-encryption
!
hostname Core1
!
enable password cisco
!
no aaa new-model
ip subnet-zero
ip routing
no ip domain-lookup
!
no file verify auto
!
spanning-tree mode pvst
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
no spanning-tree vlan 1-4094
!
vlan internal allocation policy ascending
!
!
interface Port-channel1
 switchport access vlan 35
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
!
interface Port-channel2
 switchport access vlan 35
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
!
interface GigabitEthernet0/1
 switchport access vlan 35
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 1 mode on
!
interface GigabitEthernet0/2
 switchport access vlan 35
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 1 mode on
!
interface GigabitEthernet0/9
 switchport access vlan 99
 switchport mode access
!
interface GigabitEthernet0/11
 switchport access vlan 35
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 2 mode active
!
interface GigabitEthernet0/12
 switchport access vlan 35
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 2 mode active
!
interface GigabitEthernet0/13
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 5
 switchport mode trunk
!
interface GigabitEthernet0/23
 no switchport
 ip address 10.56.0.245 255.255.255.0 -- .246 on core 2
!
interface GigabitEthernet0/24
 no switchport
 ip address 9.42.171.245 255.255.255.0 -- .246 on core 2
!
interface Vlan1
 no ip address
 shutdown
!
interface Vlan35
! -- core 1 only
 ip address 10.35.0.245 255.255.255.0
!
! interface Vlan46 -- core 2 only
! ip address 10.46.0.246 255.255.255.0
! shutdown
!
interface Vlan99
 ip address 10.99.0.245 255.255.255.0
! address .246 on core 2
!
!
ip default-gateway 9.42.171.3
ip classless
ip route 0.0.0.0 0.0.0.0 9.42.171.3
ip route 10.10.0.0 255.255.255.0 10.35.0.100
! -- points to 10.46.0.100 on core 2
ip route 10.20.0.0 255.255.255.0 10.35.0.100
! -- points to 10.46.0.100 on core 2
ip http server
ip http secure-server
!
!
!
control-plane
!
!
line con 0
line vty 0 4
 password cisco
 no login
```

```
line vty 5 15
 no login
!
!
end
```

7.8.2 Dynamic routing options OSPF/RIP

This section discusses an enhancement to the Layer 3 configuration that is shown in 7.8.1, “Layer 3 sample configuration with static routing and VRRP” on page 109. Instead of using static routes to enable the switches to know how to reach the various subnets, standard dynamic routing protocols is used. In addition to the use of the dynamic routing protocols, this configuration has a slightly different topology than the one above, introducing two additional VLANs for connections between the GbESMs and the Core switches as well as a direct cross-connection between the two GbESMs.

The use of these protocols and this more complex topology allows a Layer 3 configuration which provides High Availability without requiring the use of VRRP. The instructions that follow assume starting from the configuration that is described in 7.8.1, “Layer 3 sample configuration with static routing and VRRP” on page 109.

Merits of this configuration

This configuration is robust and suited for environments requiring High Availability. Its behavior in two key failure modes is:

- Switch failure

  In the event of a switch failure, the second switch in the BladeCenter chassis will take on the functions of the failed switch. Upstream (Core1 and Core2 in the samples) switches updates their routing tables to reflect the absence of the connection to the failed switch and sends inbound traffic to the second switch. NIC teaming is used to ensure that traffic is sent by the server blades to the second switch.

- Uplink failure

  In the event of an uplink failure to the primary switch, upstream (Core1 or Core2) switches will update their routing tables and send traffic only to the second GbESM.

  Outbound traffic will go to the primary switch and then through the crossover link to the secondary switch unless tracking and hot-standby are used to trigger a VRRP failover.

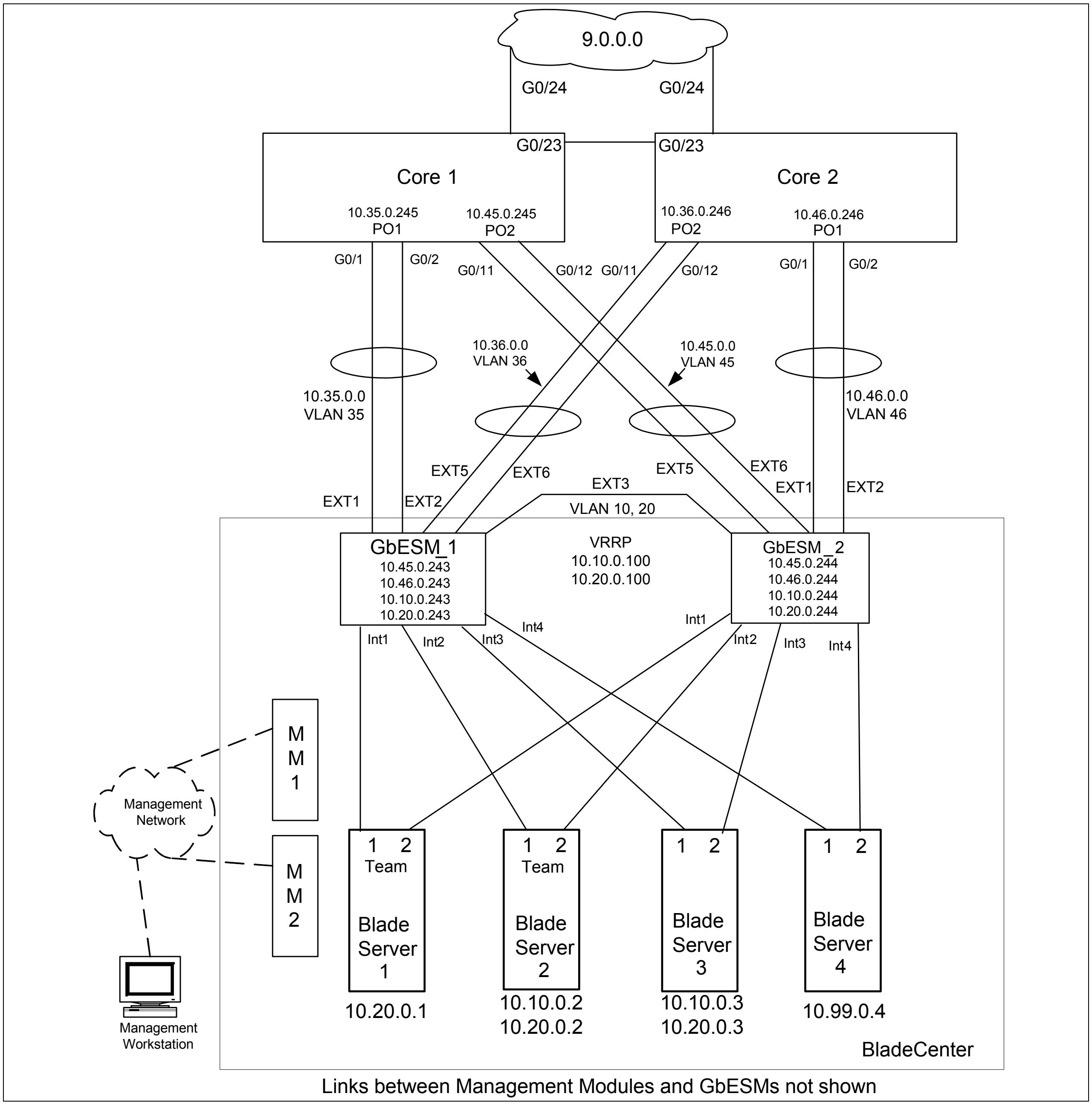
  This sample configuration does not use hot-standby. Instead, it relies on the crossover connection between the two GbESM modules to forward traffic between the two switches: the secondary switch, which connects the IBM @server BladeCenter to the outside world, and the primary switch, which continues to forward traffic to and from the server blades.

Summary of IP Addressing used in this sample

Table 7-2 summarizes the IP addresses in use. Figure 7-7 also shows these.

Table 7-2 IP Addresses for Layer 3 Sample Configuration (Dynamic routing)

Switch	VLAN 35	VLAN 36	VLAN 45	VLAN 46	VLAN 10	VLAN 20
GbESM 1	10.35.0.243	10.36.0.243	none	none	10.10.0.243	10.20.0.243
GbESM 2	none	none	10.45.0.244	10.46.0.244	10.10.0.244	10.20.0.244
Core 1	10.35.0.245	none	10.45.0.245	none	none	none
Core 2	none	10.36.0.246	none	10.46.0.246	none	none



Links between Management Modules and GbESMs not shown

Figure 7-7 Layer 3 Sample Configuration (Dynamic routing)

Configure additional VLANs and subnets for the uplink ports

Note: The configuration instructions that follow build in part on the configuration in section 7.8.1, “Layer 3 sample configuration with static routing and VRRP” on page 109

Example 7-29 and Example 7-30 on page 129 use a VLAN and associated subnet for each pair of switches which are connected. VLAN 35 connects GbESM1 (.243) with Core1 (.245), and similarly for VLANs 36, 45, and 46. The port aggregation (trunking) used in the previous samples is used here as well. The GbESMs still require default gateways on the VLANs which connect them to the Core switches. This is expected to change in future software releases.

Example 7-29 Additional VLANs and Subnets for uplink ports - GbESM

```
/* Configuration below for GbESM 1 (address .243)
/* turn layer 3 functions on - and define default gateways
/cfg/l3/frwd/on

/cfg/l3/gw 1
addr 10.35.0.245
ena
/cfg/l3/gw 2
addr 10.36.0.246
ena

/* create new vlans for uplink ports and reassign ports as needed
/cfg/l2/vlan 35/ena
/cfg/l2/vlan 36/ena

/cfg/port ext1/tag dis
/cfg/port ext2/tag dis
/cfg/port ext3/tag ena
/cfg/port ext5/tag dis
/cfg/port ext6/tag dis

/cfg/l2/vlan 35/def ext1 ext2
/cfg/l2/vlan 36/def ext5 ext6

/cfg/l2/vlan 10/add ext3
/cfg/l2/vlan 20/add ext3

   /* create VLAN interfaces

/cfg/l3/if 35
   ena
   vlan 35
   addr 10.35.0.243
   mask 255.255.255.0

/cfg/l3/if 36
   ena
   vlan 36
   addr 10.36.0.243
   mask 255.255.255.0

/* the below lines clean up unused items from the previous sample
/cfg/vlan 46/del
/cfg/l3/if 46/del
/cfg/l3/vrrp/vrid 3/del
```

```
/cfg/l3/vrrp/vrid 4/del
/cfg/l3/vrrp/hot dis
/cfg/l3/vrrp/group/dis


/* Configuration below for GbESM 2 (address .244) only

/cfg/l3/frwd/on

/cfg/l3/gw 1
addr 10.45.0.245
ena
/cfg/l3/gw 2
addr 10.46.0.246
ena
/* create new vlans for uplink ports and reassign ports as needed
/cfg/l2/vlan 45/ena
/cfg/l2/vlan 46/ena

/cfg/port ext1/tag dis
/cfg/port ext2/tag dis
/cfg/port ext3/tag ena
/cfg/port ext5/tag dis
/cfg/port ext6/tag dis

/cfg/l2/vlan 46/def ext1 ext2
/cfg/l2/vlan 45/def ext5 ext6

/cfg/l2/vlan 10/add ext3
/cfg/l2/vlan 20/add ext3

/cfg/l3/if 45
   ena
   vlan 45
   addr 10.45.0.244
   mask 255.255.255.0

/cfg/l3/if 46
   ena
   vlan 46
   addr 10.46.0.244
   mask 255.255.255.0

/* the below lines clean up unused items from the previous sample
/cfg/vlan 46/del
/cfg/l3/if 46/del
/cfg/l3/vrrp/vrid 3/del
/cfg/l3/vrrp/vrid 4/del
/cfg/l3/vrrp/hot dis
/cfg/l3/vrrp/group/dis
```

Example 7-30 Additional VLANs and Subnets - Core switches

```
! Configuration for Core 1 (address .245)
! first of all turn layer 3 functions on
   ip routing
!
! explicitly create new VLAN & remove unused - this does not appear in sh run
   vlan 45
   no vlan 46
!
   interface range Po1, g0/1-2
   switchport access vlan 35
   switchport mode access

   interface range Po2, g0/11-12
   switchport access vlan 45
   switchport mode access

   interface vlan 35
   ip address 10.35.0.245 255.255.255.0
   interface vlan 45
   ip address 10.45.0.245 255.255.255.0

! Configuration for Core 2 (address .246)
   ip routing
!
! explicitly create new VLAN & remove unused - this does not appear in sh run
   vlan 36
   no vlan 35
!
   interface range Po1, g0/1-2
   switchport access vlan 46
   switchport mode access

   interface range Po2, g0/11-12
   switchport access vlan 36
   switchport mode access

   interface vlan 36
   ip address 10.36.0.246 255.255.255.0
   interface vlan 46
   ip address 10.46.0.246
```

Configuring RIP

Routing Information Protocol (RIP) is the oldest and simplest of the dynamic routing protocols but it is nonetheless adequate for use in many BladeCenter networking configurations. The configurations in Example 7-31 and Example 7-32 use RIP version 2, which is supported by all of the devices used in our testing. Example 7-31 shows the addition of RIP to the GbESM switches. Example 7-32 shows the addition of RIP to the Core switches.

Example 7-31 RIP configuration for the GbESMs

```
/c/l3/rip/on

/c/l3/rip/if 10/enable
/c/l3/rip/if 20/enable
/c/l3/rip/if 35/enable
/* if 45 for GbESM 2
/c/l3/rip/if 36/enable
/* if 46 for GbESM 2
```

Example 7-32 RIP configuration for the Core switches

```
! first remove static routes
no ip route 10.10.0.0 255.255.255.0
no ip route 10.20.0.0 255.255.255.0
!
router rip
version 2
network 10.0.0.0
```

The redistribute static command was included in our testing to ensure that the route from our lab network to the rest of the building would be known by all switches in the test environment. It might or might not be useful in other environments.

Configuring OSPF

Open Shortest Path First (OSPF) is a more complex dynamic routing protocol. It has advantages over RIP, including the ability to support larger networks and a faster recovery time from outages. The choice of routing protocol is usually made by those individuals with responsibility for the overall network design.

These example configurations use OSPF version 2, which is supported by all of the devices that we used in our testing except the Extreme switches.

Example 7-33 shows the addition of OSPF to the GbESM switches. It is assumed that RIP is not present on the switches when the configuration of OSPF is begun. The configuration is identical on the two GbESMs except as noted.

Example 7-33 OSPF configuration for GbESMs

```
/c/l3/ospf/on
/c/l3/ospf/aindex 0
   areaid 0.0.0.9
   ena

/c/l3/ospf/if 10
   aindex 0
   ena
/c/l3/ospf/if 20
   aindex 0
   ena
```

```
/c/l3/ospf/if 35
/* -- if 45 on GbESM 2
   aindex 0
   ena

/c/l3/ospf/if 36
/* -- if 46 on GbESM 2
   aindex 0
   ena

/* also turn RIP (from previous sample) off
/cfg/l3/rip/off
```

Example 7-34 and Example 7-35 show the addition of OSPF to the Core switches. As in Example 7-33 on page 130, it is assumed that RIP is not present or has been removed.

Example 7-34   Configuration for Core switch 1

```
router ospf 9
 network 10.35.0.0 0.0.255.255 area 0.0.0.9
 network 10.45.0.0 0.0.255.255 area 0.0.0.9
! turn RIP off
no router rip
```

Example 7-35   Configuration for Core switch 2

```
router ospf 9
 network 10.36.0.0 0.0.255.255 area 0.0.0.9
 network 10.46.0.0 0.0.255.255 area 0.0.0.9
!turn rip off
no router rip
```

Verifying operation of this configuration

The following commands verify that the environment is working as advertised. As with the configuration examples, you can use the **ping** and **tracert/traceroute** commands to test end-to-end connectivity.

GbESM commands to verify RIP

The **/i/l3/route/dump** command (as shown in Example 7-36) shows the entire IP routing table on the GbESM. Routes learned from other routers are shown as `indirect rip` or `indirect ospf`. Other routes are shown as `direct` or `local`. This is a very useful command in all Layer 3 configurations. The **/i/l3/rip/dump** and **/i/l3/rip/route** commands provide more detailed information about the operation of RIP.

Example 7-36   GbESM commands to verify RIP

```
/i/l3/route/dump
Status code: * - best
    Destination        Mask         Gateway        Type      Tag    Metr If
  --------------- --------------- --------------- --------- --------- ---- --
  0.0.0.0         0.0.0.0         9.42.171.242    indirect  static         128
* 9.42.171.0      255.255.255.0   9.42.171.243    direct    fixed          128
* 9.42.171.243    255.255.255.255 9.42.171.243    local     addr           128
* 9.42.171.255    255.255.255.255 9.42.171.255    broadcast broadcast      128
* 10.10.0.0       255.255.255.0   10.10.0.243     direct    fixed          10
* 10.10.0.243     255.255.255.255 10.10.0.243     local     addr           10
* 10.10.0.255     255.255.255.255 10.10.0.255     broadcast broadcast      10
* 10.20.0.0       255.255.255.0   10.20.0.243     direct    fixed          20
* 10.20.0.243     255.255.255.255 10.20.0.243     local     addr           20
* 10.20.0.255     255.255.255.255 10.20.0.255     broadcast broadcast      20
* 10.35.0.0       255.255.255.0   10.35.0.243     direct    fixed          35
* 10.35.0.243     255.255.255.255 10.35.0.243     local     addr           35
* 10.35.0.255     255.255.255.255 10.35.0.255     broadcast broadcast      35
* 10.36.0.0       255.255.255.0   10.36.0.243     direct    fixed          36
* 10.36.0.243     255.255.255.255 10.36.0.243     local     addr           36
* 10.36.0.255     255.255.255.255 10.36.0.255     broadcast broadcast      36
* 10.45.0.0       255.255.255.0   10.10.0.244     indirect  rip          2 10
* 10.46.0.0       255.255.255.0   10.10.0.244     indirect  rip          2 10
* 10.99.0.0       255.255.255.0   10.36.0.246     indirect  rip          2 36
* 127.0.0.0       255.0.0.0       0.0.0.0         martian   martian
* 255.255.255.255 255.255.255.255 255.255.255.255 broadcast broadcast

>> GbESM_1 - IP Routing# /i/l3/rip
------------------------------------------------------------
[RIP Information Menu]
     routes   - Show RIP routes
     dump     - Show RIP user's configuration

>> GbESM_1 - RIP Information# routes
10.10.0.0/24 via 10.20.0.244 metric 2
10.20.0.0/24 via 10.10.0.244 metric 2
10.35.0.0/24 via 10.20.0.244 metric 3
10.36.0.0/24 via 10.20.0.244 metric 3
10.45.0.0/24 via 10.10.0.244 metric 2
10.46.0.0/24 via 10.10.0.244 metric 2
10.99.0.0/24 via 10.36.0.246 metric 2

>> GbESM_1 - RIP Information# dump
Enter interface number: (1-128) or 0 to show all 10

RIP USER CONFIGURATION :
   RIP on updat 30
   RIP Interface 10 : 10.10.0.243,     enabled
   version 2, listen enabled, supply enabled, default listen
   poison disabled, trigg enabled, mcast enabled, metric 1
   auth none,key none
```

Cisco commands to verify RIP

The **sh ip route** command is the CIsco equivalent of the dump command shown above. RIP routes are marked by the letter R in the left hand column. Most OSPF commands are marked by the letter O. The **show ip protocol** command (shown in Example 7-37) provides detailed information about the operation of RIP when it is running (and also on OSPF when it is running).

Example 7-37 Cisco commands to verify RIP

```
sh ip prot
*** IP Routing is NSF aware ***

Routing Protocol is "rip"
  Sending updates every 30 seconds, next due in 19 seconds
  Invalid after 180 seconds, hold down 180, flushed after 240
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Redistributing: rip
  Default version control: send version 2, receive version 2
```

```
    Interface                 Send  Recv  Triggered RIP  Key-chain
    Vlan35                    2     2
    Vlan36                    2     2
    Vlan45                    2     2
    Vlan46                    2     2
    GigabitEthernet0/23       2     2
    GigabitEthernet0/24       2     2
  Automatic network summarization is in effect
  Maximum path: 4
  Routing for Networks:
    10.0.0.0
  Routing Information Sources:
    Gateway         Distance      Last Update
    10.46.0.244          120      00:00:00
    10.36.0.243          120      00:00:02
  Distance: (default is 120)

Core2#sh ip rip database
10.0.0.0/8    auto-summary
10.10.0.0/24
    [1] via 10.46.0.244, 00:00:07, Vlan46
    [1] via 10.36.0.243, 00:00:07, Vlan36
10.20.0.0/24
    [1] via 10.46.0.244, 00:00:07, Vlan46
    [1] via 10.36.0.243, 00:00:07, Vlan36
10.35.0.0/24
    [1] via 10.36.0.243, 00:00:07, Vlan36
10.36.0.0/24    directly connected, Vlan36
10.45.0.0/24
    [1] via 10.46.0.244, 00:00:07, Vlan46
10.46.0.0/24    directly connected, Vlan46
10.99.0.0/24    directly connected, GigabitEthernet0/24
```

GbESM commands to verify OSPF

In addition to the **/i/l3/route/dump** command shown in Example 7-37 on page 132, **/i/l3/ospf/dump** provides detailed OSPF information. The **/i/l3/ospf/nbr** command lists the neighboring switches that the GbESM is connected to via OSPF. Unlike RIP, OSPF uses persistent connections with its neighbors which can be listed (see Example 7-38).

Example 7-38 GbESM commands to verify OSPF

```
/i/l3/route/du
Status code: * - best
    Destination        Mask          Gateway        Type      Tag    Metr If
  --------------- --------------- --------------- --------- --------- ---- --
  0.0.0.0         0.0.0.0         10.36.0.246     indirect  ospf         1 36
  0.0.0.0         0.0.0.0         9.42.171.242    indirect  static         128
* 9.42.171.0      255.255.255.0   9.42.171.243    direct    fixed          128
* 9.42.171.243    255.255.255.255 9.42.171.243    local     addr           128
* 9.42.171.255    255.255.255.255 9.42.171.255    broadcast broadcast      128
* 10.10.0.0       255.255.255.0   10.10.0.243     direct    fixed          10
* 10.10.0.243     255.255.255.255 10.10.0.243     local     addr           10
* 10.10.0.255     255.255.255.255 10.10.0.255     broadcast broadcast      10
* 10.20.0.0       255.255.255.0   10.20.0.243     direct    fixed          20
* 10.20.0.243     255.255.255.255 10.20.0.243     local     addr           20
* 10.20.0.255     255.255.255.255 10.20.0.255     broadcast broadcast      20
* 10.35.0.0       255.255.255.0   10.35.0.243     direct    fixed          35
* 10.35.0.243     255.255.255.255 10.35.0.243     local     addr           35
* 10.35.0.255     255.255.255.255 10.35.0.255     broadcast broadcast      35
* 10.36.0.0       255.255.255.0   10.36.0.243     direct    fixed          36
* 10.36.0.243     255.255.255.255 10.36.0.243     local     addr           36
* 10.36.0.255     255.255.255.255 10.36.0.255     broadcast broadcast      36
* 10.45.0.0       255.255.255.0   10.20.0.244     indirect  ospf         2 20
* 10.45.0.0       255.255.255.0   10.10.0.244     indirect  ospf         2 10
* 10.46.0.0       255.255.255.0   10.20.0.244     indirect  ospf         2 20
* 10.46.0.0       255.255.255.0   10.10.0.244     indirect  ospf         2 10
* 10.46.0.0       255.255.255.0   10.36.0.246     indirect  ospf         2 36
* 127.0.0.0       255.0.0.0       0.0.0.0         martian   martian
* 224.0.0.0       240.0.0.0       0.0.0.0         multicast addr
* 224.0.0.0       224.0.0.0       0.0.0.0         martian   martian
* 224.0.0.2       255.255.255.255 0.0.0.0         multicast addr
* 224.0.0.5       255.255.255.255 0.0.0.0         multicast addr
* 224.0.0.6       255.255.255.255 0.0.0.0         multicast addr
* 224.0.0.18      255.255.255.255 0.0.0.0         multicast addr
* 255.255.255.255 255.255.255.255 255.255.255.255 broadcast broadcast

>> GbESM_1 - IP Routing# /i/l3/ospf
------------------------------------------------------------
[OSPF Information Menu]
     general  - Show general information
     aindex   - Show area(s) information
     if       - Show interface(s) information
     virtual  - Show details of virtual links
     nbr      - Show neighbor(s) information
     dbase    - Database Menu
     sumaddr  - Show summary address list
     nsumadd  - Show NSSA summary address list
     routes   - Show OSPF routes
     dump     - Show OSPF information

>> GbESM_1 - OSPF Information# nbr
Intf  NeighborID       Prio  State         Address
----  ----------       ----  -----         -------
  10  9.42.171.244        1  Full           10.10.0.244
  20  9.42.171.244        1  Full           10.20.0.244
  36  10.56.0.246         1  Full           10.36.0.246

>> GbESM_1 - OSPF Information# routes
Codes: IA - OSPF inter area,
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       *  - best

*E2 0.0.0.0/0 via 10.36.0.246
    10.10.0.0/24 via 10.10.0.0
    10.20.0.0/24 via 10.20.0.0
    10.35.0.0/24 via 10.35.0.0
    10.36.0.0/24 via 10.36.0.0
*   10.45.0.0/24 via 10.20.0.244
*   10.45.0.0/24 via 10.10.0.244
*   10.46.0.0/24 via 10.20.0.244
*   10.46.0.0/24 via 10.10.0.244
*   10.46.0.0/24 via 10.36.0.246

>> GbESM_1 - OSPF Information# dump

OSPF Information:
OSPF Version 2
Router ID: 10.10.0.243
```

```
Started at 9936 and the process uptime is 11418
Area Border Router: no, AS Boundary Router: no
External LSA count 1
Number of interfaces in this router is 4
Number of virtual links in this router is 0
61 new lsa received and 21 lsa originated from this router
Total number of entries in the LSDB 9
Total neighbors are 3, of which
                                  3 are >=INIT state,
                                  3 are >=EXCH state,
                                  3 are =FULL state
Number of areas is 1, of which 1-transit 0-nssa
   Area Id : 0.0.0.9
   Authentication : none
   Import ASExtern : yes
   Number of times SPF ran : 11
   Area Border Router count : 0
   AS Boundary Router count : 1
   LSA count : 8
   Summary : noSummary
------------------------------------------------------------------

OSPF Neighbors:
Intf  NeighborID       Prio  State         Address
----  ----------       ----  -----         -------
  10  9.42.171.244        1  Full           10.10.0.244
  20  9.42.171.244        1  Full           10.20.0.244
  36  10.56.0.246         1  Full           10.36.0.246

OSPF LS Database:
   OSPF LSDB breakdown for router with ID (10.10.0.243)
  Area IDRouterNetworkSum-NetASBRNSSASubtotal
  0.0.0.9440 0  0  8
  AS External      1
  Total4  4  0  0  0  9

               AS External LSAs (Area 0.0.0.9)

  Link ID         ADV Router      Options Age   Seq#         Checksum
  0.0.0.0         10.56.0.246     0x20    447   0x80000007   0xD991

               Router LSAs (Area 0.0.0.9)

  Link ID         ADV Router      Options Age   Seq#         Checksum
  10.56.0.246     10.56.0.246     0x22    1441  0x80000010   0x23D5
  9.42.171.243    9.42.171.243    0x2     1658  0x80000012   0xAC14
  9.42.171.244    9.42.171.244    0x2     1474  0x80000017   0x6D2D
  10.10.0.243     10.10.0.243     0x2     1440  0x80000003   0x491C

               Network LSAs (Area 0.0.0.9)

  Link ID         ADV Router      Options Age   Seq#         Checksum
  10.36.0.246     10.56.0.246     0x22    1440  0x80000008   0xBDD5
  10.20.0.244     9.42.171.244    0x2     1484  0x80000007   0xC7C9
  10.46.0.244     9.42.171.244    0x2     1415  0x80000008   0xE064
  10.10.0.244     9.42.171.244    0x2     1484  0x80000007   0x405B
```

Cisco commands to verify OSPF

The following commands are in addition to the Cisco commands used for RIP (see Example 7-39):

sh ip ospf neighbor Lists the OSPF neighbors of the current switch.

sh ip ospf database Provides a dump of the OSPF topology database, which includes all switches in the same area as the one where the command is issued.

Example 7-39 Cisco commands to verify OSPF

```
sh ip route

Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is 9.42.171.3 to network 0.0.0.0

     9.0.0.0/24 is subnetted, 1 subnets
C       9.42.171.0 is directly connected, GigabitEthernet0/24
     10.0.0.0/24 is subnetted, 7 subnets
O       10.10.0.0 [110/2] via 10.46.0.244, 00:25:13, Vlan46
                  [110/2] via 10.36.0.243, 00:25:13, Vlan36
O       10.20.0.0 [110/2] via 10.46.0.244, 00:25:13, Vlan46
                  [110/2] via 10.36.0.243, 00:25:13, Vlan36
C       10.46.0.0 is directly connected, Vlan46
O       10.45.0.0 [110/2] via 10.46.0.244, 00:25:13, Vlan46
O       10.35.0.0 [110/2] via 10.36.0.243, 00:25:13, Vlan36
C       10.36.0.0 is directly connected, Vlan36
C       10.99.0.0 is directly connected, GigabitEthernet0/24
S*   0.0.0.0/0 [1/0] via 9.42.171.3


Core2#sh ip ospf neighbor

Neighbor ID     Pri   State           Dead Time   Address         Interface
9.42.171.244      1   FULL/DR         00:00:31    10.46.0.244     Vlan46
10.10.0.243       1   FULL/BDR        00:00:39    10.36.0.243     Vlan36
Core2#sh ip ospf database

            OSPF Router with ID (10.56.0.246) (Process ID 9)

      Router Link States (Area 0.0.0.9)

Link ID         ADV Router      Age         Seq#       Checksum Link count
9.42.171.243    9.42.171.243    1750        0x80000012 0x00AC14 4
9.42.171.244    9.42.171.244    1567        0x80000017 0x006D2D 4
10.10.0.243     10.10.0.243     1537        0x80000003 0x00491C 4
10.56.0.246     10.56.0.246     1532        0x80000010 0x0023D5 2

      Net Link States (Area 0.0.0.9)

Link ID         ADV Router      Age         Seq#       Checksum
10.10.0.244     9.42.171.244    1577        0x80000007 0x00405B
10.20.0.244     9.42.171.244    1577        0x80000007 0x00C7C9
```

```
10.36.0.246     10.56.0.246     1532        0x80000008 0x00BDD5
10.46.0.244     9.42.171.244    1507        0x80000008 0x00E064

      Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
0.0.0.0         10.56.0.246     517         0x80000007 0x00D991 9
```

Complete configuration snapshots

Configurations shown in Example 7-40 and Example 7-41 on page 140 include details for OSPF and RIP. In general they should not be run at the same time. One configuration for GbESMs and one for Core switches is shown with notes when the second switch of the same type has a different configuration.

Example 7-40 Layer 3 GbESM configuration with dynamic routing

```
>> GbESM_1 - Main# /c/d
script start "Layer 2-3 Gigabit Ethernet Switch Module for IBM eServer BladeCenter" 4
/**** DO NOT EDIT THIS LINE!
/* Configuration dump taken  0:04:01 Thu Jan  1, 2070
/* Version 1.0.1.6,  Base MAC address 00:11:f9:36:b7:00
/* GbESM_1
/c/sys
        hprompt ena
/c/sys/access/user/uid 1
        name "USERID"
        pswd "177d191d054908081634f2f6c6f9abfa98b78d75d776b757b047e9f5e089cb8c"
        ena
        cos admin
/c/sys/ssnmp
        name "GbESM_1"
/c/port INT1
        pvid 20
/c/port INT2
        pvid 10
/c/port INT3
        pvid 10
/c/port EXT1
        pvid 35
/c/port EXT2
        pvid 35
/c/port EXT3
        tag ena
        pvid 10
/c/port EXT5
        pvid 36
/c/port EXT6
        pvid 36
/c/l2/vlan 1
        def INT1 INT2 INT3 INT4 INT5 INT6 INT7 INT8 INT9 INT10 INT11 INT12 INT13 INT14 EXT4
/c/l2/vlan 10
        ena
        name "VLAN_Green"
        def INT2 INT3 EXT3
/* above excludes INT3 on GbESM2
/c/l2/vlan 20
        ena
        name "VLAN_Red"
        def INT1 INT2 EXT3
```

```
/* above also INT3 on GbESM 2
/c/l2/vlan 35
/* use vlan 45 on GbESM 2
       ena
        name "VLAN 35"
        def EXT1 EXT2
/c/l2/vlan 36
/* use vlan 46 on GbESM 2
        ena
        name "VLAN 36"
        def EXT5 EXT6

/c/l2/stg 1/off
/c/l2/stg 1/clear
/c/l2/stg 1/add 1 10 20 35 36 50
/c/l2/trunk 1
        ena
        add EXT1
        add EXT2
/c/l2/trunk 2
        ena
        add EXT5
        add EXT6

/c/l3/if 10
        ena
        addr 10.10.0.243
/* -- 10.10.0.244 on GbESM 2
        mask 255.255.255.0
        broad 10.10.0.255
        vlan 10
/c/l3/if 20
        ena
        addr 10.20.0.243
/* -- 10.20.0.244 on GbESM 2
        mask 255.255.255.0
        broad 10.20.0.255
        vlan 20
/c/l3/if 35
/* -- this section uses if 45, vlan 45, address .244 on GbESM 2
        ena
        addr 10.35.0.243
        mask 255.255.255.0
        broad 10.35.0.255
        vlan 35
/c/l3/if 36
/* -- this section uses if 46, vlan 46, address .244 on GbESM 2
        ena
        addr 10.36.0.243
        mask 255.255.255.0
        broad 10.36.0.255
        vlan 36
/c/l3/vrrp/on
/c/l3/vrrp/vr 1
        ena
        vrid 1
        if 10
        prio 101
/* -- above only on GbESM 1, defaults to 100 on GbESM 2
        addr 10.10.0.100
```

```
/c/l3/vrrp/vr 2
        ena
        vrid 2
        if 20
        prio 101
/* -- above only on GbESM 1
        addr 10.20.0.100

/* the section below is for RIP - remove it for OSPF
/* for GbESM2 use if 45 and 46 instead of 35 and 36 in the section below

/c/l3/rip/updat 30/on
/c/l3/rip/if 10/ena/supply e/listen e/default listen/version 2
/c/l3/rip/if 10/poison d/trigg e/metric 1/mcast e
/c/l3/rip/if 10/auth none
/c/l3/rip/if 20/ena/supply e/listen e/default listen/version 2
/c/l3/rip/if 20/poison d/trigg e/metric 1/mcast e
/c/l3/rip/if 20/auth none
/c/l3/rip/if 35/ena/supply e/listen e/default listen/version 2
/c/l3/rip/if 35/poison d/trigg e/metric 1/mcast e
/c/l3/rip/if 35/auth none
/c/l3/rip/if 36/ena/supply e/listen e/default listen/version 2
/c/l3/rip/if 36/poison d/trigg e/metric 1/mcast e
/c/l3/rip/if 36/auth none
/
/* the section below is for OSPF - remove it for RIP
/c/l3/ospf/on
/c/l3/ospf/aindex 1
        ena
        areaid 0.0.0.9
        type transit
        metric 1
        auth none
        spf 10
/c/l3/ospf/if 10
        ena
        aindex 1
        prio 1
        cost 1
        hello 10
        dead 40
        trans 5
        retra 5
/c/l3/ospf/if 20
        ena
        aindex 1
        prio 1
        cost 1
        hello 10
        dead 40
        trans 5
        retra 5
/c/l3/ospf/if 35
/* use if 45 instead of 35 for GbESM 2
        ena
        aindex 1
        prio 1
        cost 1
        hello 10
        dead 40
        trans 5
        retra 5
/c/l3/ospf/if 36
/* use if 46 instead of 36 for GbESM 2
        ena
        aindex 1
        prio 1
        cost 1
        hello 10
        dead 40
        trans 5
        retra 5
script end  /**** DO NOT EDIT THIS LINE!
```

Example 7-41 Layer 3 Core switch configuration with dynamic routing

```
Core1#sh run

! note that VLANs must be created with the Vlan <x> command in “config t” mode
! these statements are not displayed by the “sh run” command

Building configuration...

Current configuration : 2935 bytes
!
version 12.2
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname Core1
!
enable password cisco
!
no aaa new-model
ip subnet-zero
ip routing
no ip domain-lookup
!
no file verify auto
!
spanning-tree mode pvst
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
no spanning-tree vlan 1-4094
!
vlan internal allocation policy ascending
!
!
interface Port-channel1
 switchport access vlan 35
! -- vlan 36 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
!
interface Port-channel2
```

```
 switchport access vlan 45
! -- vlan 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
!
interface GigabitEthernet0/1
 switchport access vlan 35
! -- 36 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 1 mode on
!
interface GigabitEthernet0/2
 switchport access vlan 35
! -- 36 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 1 mode on
!
interface GigabitEthernet0/9
 switchport access vlan 99
 switchport mode access
!
interface GigabitEthernet0/11
 switchport access vlan 45
! -- 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 2 mode on
!
interface GigabitEthernet0/12
 switchport access vlan 45
! -- 46 on core 2
 switchport trunk encapsulation dot1q
 switchport mode access
 switchport nonegotiate
 channel-group 2 mode on
!
!back door to ensure a working address to get in to the switch
interface GigabitEthernet0/23
 no switchport
 ip address 10.56.0.245 255.255.255.0
! -- .246 on core 2
!
interface GigabitEthernet0/24
 no switchport
 ip address 9.42.171.245 255.255.255.0
! -- .246 on core 2

!
interface Vlan1
 no ip address
 shutdown
!
! core 2 uses vlans 36 and 46 and addresses ending in .246
!
```

```
interface Vlan35
 ip address 10.35.0.245 255.255.255.0
!
interface Vlan45
 ip address 10.45.0.245 255.255.255.0
!
interface Vlan99
 ip address 10.99.0.245 255.255.255.0
!

! the below is for RIP - remove it for OSPF
router rip
 redistribute static
 network 10.0.0.0

! the below is for OSPF - remove it for RIP
! core 2 uses 10.36.0.0 and 10.46.0.0
router ospf 9
 network 10.35.0.0 0.0.255.255 area 0.0.0.9
network 10.45.0.0 0.0.255.255 area 0.0.0.9


! the below is used for OSPF and RIP

ip default-gateway 9.42.171.3
ip classless
ip route 0.0.0.0 0.0.0.0 9.42.171.3
ip http server
ip http secure-server
!
!
!
control-plane
!
!
line con 0
line vty 0 4
 password cisco
 no login
line vty 5 15
 no login
!
!
end
```

7.9 Configuration for Extreme switches

The configuration in Example 7-42 is from one of a pair of Extreme switches which we tested in place of the Cisco Core switches. The Extreme switches had a subset of the functionality of the GbESM switch modules, so our testing was limited. We were able to successfully test the following:

- VLAN tagging using 802.1Q
- Link aggregation (trunking) with static trunks; the Extreme switch did not support LACP.
- RIP

Testing of the different variations of Spanning Tree Protocol was only partially successful. This is due to some limitations of the Extreme switch's functionality as well as certain proprietary enhancements to the standards which the GbESM did not support. We were able to configure OSPF but not test it because an additional license which we did not have was required to enable this function.

Note that certain default configuration text was removed from the below due to extreme length. This test was present when we reset the switch to its factory defaults.

Example 7-42 Extreme switch configuration

```
# Full Detail Configuration

#
# Summit400-48t Configuration generated Tue Jun 14 05:33:14 2005
# Software Version 7.2e.1 (Build 10) [non-ssh] by Release_Master on 03/26/04 18:29:56

configure sys-recovery-level none
enable system-watchdog
configure reboot-loop-protection threshold 0
configure vlan default delete ports all
create vlan "VLAN10"
create vlan "VLAN20"
create vlan "VLAN99"
#
# Config information for VLAN Default.
configure vlan "Default" tag 1     # VLAN-ID=0x1  Global Tag 1
configure stpd s0 add vlan "Default"
configure vlan "Default" qosprofile "QP1"
configure vlan "Default" ipaddress 9.42.171.98 255.255.255.0
configure vlan "Default" add port 1 untagged
configure vlan "Default" add port 3 untagged
configure vlan "Default" add port 4 untagged
configure vlan "Default" add port 5 untagged
configure vlan "Default" add port 6 untagged
configure vlan "Default" add port 7 untagged
configure vlan "Default" add port 8 untagged
configure vlan "Default" add port 9 untagged
configure vlan "Default" add port 10 untagged
configure vlan "Default" add port 11 untagged
# similar lines for ports 13-50 omitted
#
# Config information for VLAN Mgmt.
# No IP address is configured for VLAN Mgmt.
#
# Config information for VLAN VLAN10.
configure vlan "VLAN10" tag 10     # VLAN-ID=0xa  Global Tag 4
configure stpd s0 add vlan "VLAN10"
configure vlan "VLAN10" qosprofile "QP1"
configure vlan "VLAN10" ipaddress 10.10.0.247 255.255.255.0
configure vlan "VLAN10" add port 1 tagged
configure vlan "VLAN10" add port 11 tagged
configure vlan "VLAN10" add port 23 tagged
#
# Config information for VLAN VLAN20.
configure vlan "VLAN20" tag 20     # VLAN-ID=0x14  Global Tag 5
configure stpd s0 add vlan "VLAN20"
configure vlan "VLAN20" qosprofile "QP1"
configure vlan "VLAN20" ipaddress 10.20.0.247 255.255.255.0
configure vlan "VLAN20" add port 1 tagged
configure vlan "VLAN20" add port 11 tagged
configure vlan "VLAN20" add port 23 tagged
#
# Config information for VLAN VLAN99.
configure vlan "VLAN99" tag 99     # VLAN-ID=0x63  Global Tag 6
configure stpd s0 add vlan "VLAN99"
configure vlan "VLAN99" qosprofile "QP1"
configure vlan "VLAN99" ipaddress 10.99.0.247 255.255.255.0
configure vlan "VLAN99" add port 1 tagged
configure vlan "VLAN99" add port 11 tagged
configure vlan "VLAN99" add port 23 tagged

# Boot information
use image primary

#Configuration Information
use configuration primary
delete account user
configure account admin encrypted
452eN2$1bS/.EI6vFqiY/TcN8wIo1
452eN2$1bS/.EI6vFqiY/TcN8wIo1
create account user "user" encrypted "yN/eN2$zxKAyFhkr/1cfugaePE/f0"
enable telnet access-profile none port 23
#
# Banner Configuration
#
# omitted
# SNMP Configuration
# omitted

# Load Sharing Configuration -  create two static multi-link trunks each w/ 2 ports
enable sharing 1 grouping 1,2
enable sharing 11 grouping 11,12

# Ports Configuration
# omitted

# Spanning tree information
configure stpd s0 tag 10
configure stpd s0 mode dot1d # dot1w is also available for rapid STP
configure stpd s0 port link-type broadcast 1
# identical lines for remaining ports omitted.
enable stpd s0

# MAC FDB configuration and static entries
configure fdb agingtime 300

configure ipfdb agingtime 0
```

```
# -- IP Interface[0] = "Default"
enable ipforwarding vlan "Default"
disable ipforwarding broadcast vlan "Default"
disable ipforwarding ignore-broadcast vlan "Default"
disable isq vlan "Default"
disable irdp vlan "Default"
enable icmp unreachable vlan "Default"
enable icmp redirects vlan "Default"
enable icmp port-unreachables vlan "Default"
enable icmp time-exceeded vlan "Default"
enable icmp parameter-problem vlan "Default"
enable icmp timestamp vlan "Default"
enable icmp address-mask vlan "Default"
configure ip-mtu 1500 vlan "Default"
unconfigure vlan "MacVlanDiscover" ipaddress
unconfigure vlan "Mgmt" ipaddress

# -- IP Interface[1] = "VLAN10"
enable ipforwarding vlan "VLAN10"
disable ipforwarding broadcast vlan "VLAN10"
disable ipforwarding ignore-broadcast vlan "VLAN10"
disable isq vlan "VLAN10"
disable irdp vlan "VLAN10"
enable icmp unreachable vlan "VLAN10"
enable icmp redirects vlan "VLAN10"
enable icmp port-unreachables vlan "VLAN10"
enable icmp time-exceeded vlan "VLAN10"
enable icmp parameter-problem vlan "VLAN10"
enable icmp timestamp vlan "VLAN10"
enable icmp address-mask vlan "VLAN10"
configure ip-mtu 1500 vlan "VLAN10"

# -- IP Interface[2] = "VLAN20"
enable ipforwarding vlan "VLAN20"
disable ipforwarding broadcast vlan "VLAN20"
disable ipforwarding ignore-broadcast vlan "VLAN20"
disable isq vlan "VLAN20"
disable irdp vlan "VLAN20"
enable icmp unreachable vlan "VLAN20"
enable icmp redirects vlan "VLAN20"
enable icmp port-unreachables vlan "VLAN20"
enable icmp time-exceeded vlan "VLAN20"
enable icmp parameter-problem vlan "VLAN20"
enable icmp timestamp vlan "VLAN20"
enable icmp address-mask vlan "VLAN20"
configure ip-mtu 1500 vlan "VLAN20"

# -- IP Interface[3] = "VLAN99"
enable ipforwarding vlan "VLAN99"
disable ipforwarding broadcast vlan "VLAN99"
disable ipforwarding ignore-broadcast vlan "VLAN99"
disable isq vlan "VLAN99"
disable irdp vlan "VLAN99"
enable icmp unreachable vlan "VLAN99"
enable icmp redirects vlan "VLAN99"
enable icmp port-unreachables vlan "VLAN99"
enable icmp time-exceeded vlan "VLAN99"
enable icmp parameter-problem vlan "VLAN99"
enable icmp timestamp vlan "VLAN99"
enable icmp address-mask vlan "VLAN99"
```

```
configure ip-mtu 1500 vlan "VLAN99"

# Global IP settings.
configure irdp 450 600 1800 0
configure irdp broadcast
disable icmp useredirects
disable iproute sharing
configure ipfdb route-add clear-all
disable bootprelay
configure ip-down-vlan-action forward
#
# IP ARP Configuration

configure iparp timeout 20
configure iparp max-entries 4096
configure iparp max-pending-entries 256
enable iparp checking
enable iparp refresh
#
# IP Route Configuration
configure iproute add default 9.42.171.3 1
# Multicast configuration
# omitted

# RIP interface configuration
configure rip delete vlan "Default"
configure rip txmode v2only vlan "Default"
configure rip rxmode any vlan "Default"
configure rip vlan "Default" cost 1
configure rip vlan "Default" trusted-gateway None
configure rip vlan "Default" import-filter None
configure rip vlan "Default" export-filter None
configure rip add vlan "VLAN99"
configure rip txmode v2only vlan "VLAN99"
configure rip rxmode any vlan "VLAN99"
configure rip vlan "VLAN99" cost 1
configure rip vlan "VLAN99" trusted-gateway None
configure rip vlan "VLAN99" import-filter None
configure rip vlan "VLAN99" export-filter None
configure rip add vlan "VLAN20"
configure rip txmode v2only vlan "VLAN20"
configure rip rxmode any vlan "VLAN20"
configure rip vlan "VLAN20" cost 1
configure rip vlan "VLAN20" trusted-gateway None
configure rip vlan "VLAN20" import-filter None
configure rip vlan "VLAN20" export-filter None
configure rip add vlan "VLAN10"
configure rip txmode v2only vlan "VLAN10"
configure rip rxmode any vlan "VLAN10"
configure rip vlan "VLAN10" cost 1
configure rip vlan "VLAN10" trusted-gateway None
configure rip vlan "VLAN10" import-filter None
configure rip vlan "VLAN10" export-filter None
# RIP global parameter configuration
disable rip aggregation
enable rip splithorizon
enable rip poisonreverse
enable rip triggerupdate
disable rip export static
disable rip export ospf-intra
disable rip export ospf-inter
disable rip export ospf-extern1
disable rip export ospf-extern2
disable rip export direct
disable rip originate-default
configure rip updatetime 30
configure rip routetimeout 180
configure rip garbagetime 120
# RIP Global enable/disable state
enable rip

#
# PIM Router Configuration
#
disable pim
# remaining details omitted

# Ospf Area Configuration
create ospf area 0.0.0.9
configure ospf area 0.0.0.9 interarea-filter "None"
configure ospf area 0.0.0.9 external-filter "None"

# Ospf Range Configuration

# Interface Configuration
configure ospf vlan "Default" area 0.0.0.9
configure ospf vlan "Default" timer 5 1 10 40
configure ospf vlan "Default" authentication none
configure ospf vlan "VLAN99" area 0.0.0.9
configure ospf vlan "VLAN99" timer 5 1 10 40
configure ospf vlan "VLAN99" authentication none
configure ospf vlan "VLAN20" area 0.0.0.9
configure ospf vlan "VLAN20" timer 5 1 10 40
configure ospf vlan "VLAN20" authentication none
configure ospf vlan "VLAN10" area 0.0.0.9
configure ospf vlan "VLAN10" timer 5 1 10 40
configure ospf vlan "VLAN10" authentication none

# Virtual Link Configuration

# Ospf ASE Summary Configuration

# OSPF Router Configuration
configure ospf lsa-batch-interval 30
configure ospf metric-table 10M 10 100M 5 1G 4 10G 2
configure ospf spf-hold-time 3
enable ospf capability opaque-lsa
configure ospf ase-limit 0 timeout 0

disable ospf export static
disable ospf export direct
disable ospf export rip

# VRRP Configuration

# EAPS configuration
disable eaps
configure eaps fast-convergence off

# EAPS shared port configuration
# SNTP client configuration
# omitted

# Mac Vlan Configurations
#
# Access-mask Configuration
#
# Access-list Configuration
#
# Rate-limit Configuration

#
# System Dump Configuration
#

## SNMPV3 EngineID Configuration
#
## SNMPV3 USM Users Configuration
#
#
# SNMPV3 MIB Views Configuration
#
#
# SNMPV3 VACM Access Configuration
#
#
# SNMPV3 USM Groups Configuration
#
#
# SNMPV3 Community Table Configuration
#
#
# SNMPV3 Target Addr Configuration
#
#
# SNMPV3 Target Params Configuration
#
#
# SNMPV3 Notify Configuration
#
#
# SNMPV3 Notify Filter Profile Configuration
#
#
# SNMPV3 Notify Filter Configuration
#



# System-wide Debug Configuration
#No System-wide debug tracing configured

#Vlan Based Debug Configuration
#
#No Vlan-based debug-tracing configured

#Port Based Debug Configuration
#
#No Port based debug-tracing configured
```

```
# Network Login Configuration
configure netlogin base-url "network-access.net"
configure netlogin redirect-page "http://www.extremenetworks.com"
enable netlogin logout-privilege
disable netlogin Session-Refresh  3
enable netlogin web-based
enable netlogin dot1x

# Event Management System Configuration

# Event Management System Log Filter Configuration

# Event Management System Log Target Configuration
disable syslog

configure log target nvram filter "DefaultFilter" severity warning
configure log target nvram match ""
configure log target nvram format priority off date mm-dd-yyyy time hundredths host-name
off tag-name off tag-id off sequence-number off severity on event-name condition
process-name off process-id off source-function off source-line off
enable log target nvram

configure log target memory-buffer number-of-messages 1000
configure log target memory-buffer filter "DefaultFilter" severity debug-data
configure log target memory-buffer match ""
configure log target memory-buffer format priority off date mm-dd-yyyy time hundredths
host-name off tag-name off tag-id off sequence-number off severity on event-name condition
process-name off process-id off source-function off source-line off
enable log target memory-buffer

configure log target console-display filter "DefaultFilter" severity info
configure log target console-display match ""
configure log target console-display format priority off date mm-dd-yyyy time hundredths
host-name off tag-name off tag-id off sequence-number off severity on event-name condition
process-name off process-id off source-function off source-line off
disable log target console-display

# cpu denial-of-service protection configuration
disable cpu-dos-protect
# remainder omitted

#
# End of configuration file for "Summit400-48t".
#
```

8

Serial over LAN feature description and configuration

This chapter provides a brief introduction to the Serial over LAN (SOL) feature for the IBM @server BladeCenter. It also discusses the configuration of the Nortel Networks L2/3 GbESM along with rules to follow for establishing an SOL connection.

8.1 SOL overview

SOL is accomplished in the following way: Serial data that flows to and from the blade server(s) COM port is routed through the network infrastructure of the BladeCenter chassis. This network infrastructure includes the BladeCenter Management Module, the Nortel Networks L2/3 GbESM, and the onboard network adapter of the blade server. In addition, the blade server's integrated system management processor also assists in handling the serial data to and from the COM port.

The BladeCenter Management Module acts as a proxy during a SOL connection from a workstation in the network and the blade server in the BladeCenter chassis. First, the workstation uses telnet to establish a link with the Management Module. When the link is established through telnet, the workstation can then pass a command to the Management Module to interact with the serial port of any blade server in that BladeCenter chassis.

Note: The traffic path of SOL along the network is through the Management Module. It is not possible to establish a SOL session to a blade server by going through the external ports of the switch.

8.2 General rules to establish an SOL connection

To initiate an SOL connection to a blade server, you must first establish a telnet session to the BladeCenter Management Module where the blade server is located. When the telnet session has been established, you can initiate a remote console SOL session to any blade server using the CLI of the Management Module. General rules for establishing this connection are:

- You can establish as many as 20 separate telnet sessions to one BladeCenter Management Module. This gives the ability to have one SOL session active on all 14 blade servers in the BladeCenter chassis and also have six additional CLI sessions for the Management Module itself.
- If security is a concern, Secure Shell (SSH) sessions are also available so users can establish secured telnet CLI sessions to the Management Module prior to starting an SOL console redirect session to the blade server.
- Telnet sessions with the Management Module have a default timeout value of 120 seconds. If there is no telnet or SOL traffic within the timeout interval, the telnet session itself terminates.
- SOL uses the first network interface (Planar Ethernet 1) of the blade server to communicate. When this interface attempts to boot through PXE or DHCP, the network interface is reset, causing the SOL connection to drop. If booting through PXE or DHCP is required, it is recommended that the second network interface of the blade server is used (Planar Ethernet 2). This can be set in the blade server's BIOS settings.
- JS20 model blade servers require SOL for video purposes during the operating system install at least. After the operating system is installed the user can use other methods over the blade server's Ethernet interface(s) for administration (SSH for example). If BOOTP is needed for JS20, again, the second network interface is recommended.

There are no specific rules that apply to the Nortel Networks L2/3 GbESM itself. The switch by default reserves VLAN 4095 for SOL traffic.

8.3 Configuring SOL for use with the Nortel GbESM

The latest updates for the following components should be obtained prior to setting up SOL:

- BladeCenter Management Module
- Blade Server BIOS
- Blade Server Diagnostics
- Blade Server system integrated system management processor (ISMP)
- Blade Server Broadcom Ethernet firmware
- Blade Server Broadcom Ethernet device driver
- Nortel Networks L2/3 GbESM

As long as the SOL VLAN ID value in the Management Module Web interface is set to 4095, there are no other specific configuration steps for the Nortel Networks L2/3 GbESM, because the Nortel Networks L2/3 GbESM reserves VLAN 4095 for SOL. All other configuration steps involve either the Management Module or the blade server. To properly set up and configure both the Management Module or blade server, refer to the following Web site:

http://www-307.ibm.com/pc/support/site.wss/document.do?Indocid=MIGR-54666

Also of note here, you do not need to modify the BSMP IP Address Range field. This value is the base IP address for blade servers in the BladeCenter (the last octet is incremented for each blade server in the chassis). The IP address that SOL uses to communicate with the blade system management processor (BSMP) of each blade server is based on the IP address set in this field.

> **Important:** It is not recommended that you use your production subnet for the BSMP IP Address Range field. These addresses are only used by the Management Module and the blade server's ISMP for SOL translation. You cannot telnet into them or use them any other way. The Management Module cannot be bypassed when trying to open a SOL session to a blade server.

8.4 SOL use during Nortel Networks L2/3 GbESM experiments

SOL was enabled on an HS40 blade server during the experiments that we used for this paper. The blade server was set up according to the following document:

http://www-307.ibm.com/pc/support/site.wss/document.do?Indocid=MIGR-54666

The default timeout value of telnet was changed in the Management Module. This allowed for the SOL session to stay up continually for monitoring the blade server. This value cannot be set using the Management Module Web interface. To reset the timeout of telnet for the Management Module, do the following:

1. Telnet to the Management Module and login with USERID and PASSW0RD.
2. At the <system> prompt type `env -T system:mm[1]`.
3. Type `telnetcfg` to display the current setting.
4. To change the value to the desired number of seconds (4000 for example, which is just over an hour), type `telnetcfg -t 4000`.
5. Type `env` to return to the <system> prompt.

The operating system of choice for this paper was Windows 2003 Standard Edition. The blade server was located in slot 5 of the BladeCenter chassis. The slot number is used for when you are attempting to start a SOL session from the Management Module. To administer the Windows 2003 blade server through SOL:

1. Telnet to the Management Module and login with USERID and PASSW0RD.
2. At the <system> prompt type `console -T blade[5]`.
3. At the <SAC> prompt type `?` for a help menu.

Commands at the <SAC> prompt are for administration of the blade server itself. The blade server can be restarted, services can be started or shutdown, IP addresses can be changed, the kernel log can be dumped, and so forth. To see a normal Windows command prompt, do the following:

1. At the <SAC> prompt type `cmd` to make a new command channel and note the name that is shown.
2. Type `ch` and look for your channel name in the list and note the number under the `#` column.
3. Type `ch -si 1` to open the channel command prompt (1 is an example, but if the above steps were followed exactly it should be the number for your channel).
4. Login to Windows as you normally would.
5. To return to the <SAC> prompt at anytime, type `exit`.

The Linux and AIX® the SOL sessions look the same as when using SSH to administer the blade server.

9

Nortel Networks Layer 2/3 GbE Switch Module troubleshooting

In this chapter, we discuss troubleshooting techniques that you can use in support of the Nortel Networks L2/3 GbESM.

9.1 Basic rules and unique symptoms

Before going into detail about troubleshooting, it is important to first discuss certain common rules and symptoms for this environment. Certain interactions within the IBM @server BladeCenter between the Nortel Networks L2/3 GbESM and the Management Module require that certain important rules are followed. Failure to follow these rules can produce unexpected results when deploying the IBM @server BladeCenter containing an Nortel Networks L2/3 GbESM. In this section, we summarize some of these rules and the consequences of not adhering to them. We also discuss some symptoms that you might see and possible solutions for these.

9.1.1 Basic rules

Review the following rules:

1. Do *not* attach cables to the Nortel Networks L2/3 GbESM until *both* sides of the connection are configured.
   - Symptoms: No upstream connectivity, upstream network failure from spanning tree loop.
   - Solution: Keep cables disconnected or ports shut down until properly configured on *both* sides of the connection. This is an important best practice for *any* switch-to-switch connections in a production network, not just between the Nortel Networks L2/3 GbESM and its upstream connections.
2. Do *not* put Blade servers on the subnet that the Nortel Networks L2/3 GbESM uses for its management VLAN interface.
   - Symptoms: Unable to reach the switch from server blades or from external devices attached to the Management Module port.
   - Solution: Use a separate subnets for the management module.

   > **Note:** It is not possible to put external ports on VLAN 4095, which is dedicated to the Management Module. It *is* essential that the PVID of ports INT1-14 be a VLAN other than 4095 and that ports INT1-14 are each members of at least one such data-bearing VLAN

3. Make sure the Nortel Networks L2/3 GbESM firmware code is upgraded to the most current version of the Alteon OS. You can find the latest code for the Nortel Networks L2/3 GbESM at:

   `http://www.ibm.com/support`

   When you go to this Web site, do the following:

   a. Under Support topics, click **Downloads and drivers**.

   b. Under Category, select **BladeCenter (Blades)**.

   c. Click BladeCenter chassis **Hardware only**.

   d. Click **Firmware**.

   e. Scroll to find the desired *Nortel Networks L2/3 GbESM Firmware - IBM @server BladeCenter*.

9.2 Nortel Networks L2/3 GbESM troubleshooting methodology

In this section, we discuss general troubleshooting techniques and offer options for getting started.

9.2.1 General comments on troubleshooting

Because of the highly integrated nature of the Nortel Networks L2/3 GbESM within the IBM @server BladeCenter, it is usually necessary to engage several teams for anything beyond basic hardware troubleshooting. Experience has shown that the greater the communication between administrative groups, the more likely an issue can be resolved sooner rather than later.

Where do you start when troubleshooting? How do you determine whether it is a hardware or configuration issue and not a software bug?

The only true answer is *experience*.

Rather than attempting to list step-by-step procedures for every possible troubleshooting issue (which could fill several volumes and still not be complete), this section offers information about what to gather and commands that can be useful. The information here assumes that the person doing the troubleshooting has experience in such matters.

9.3 Systematic approach

Figure 9-1 provides a basic flow diagram on how to approach troubleshooting Nortel Networks L2/3 GbESM.

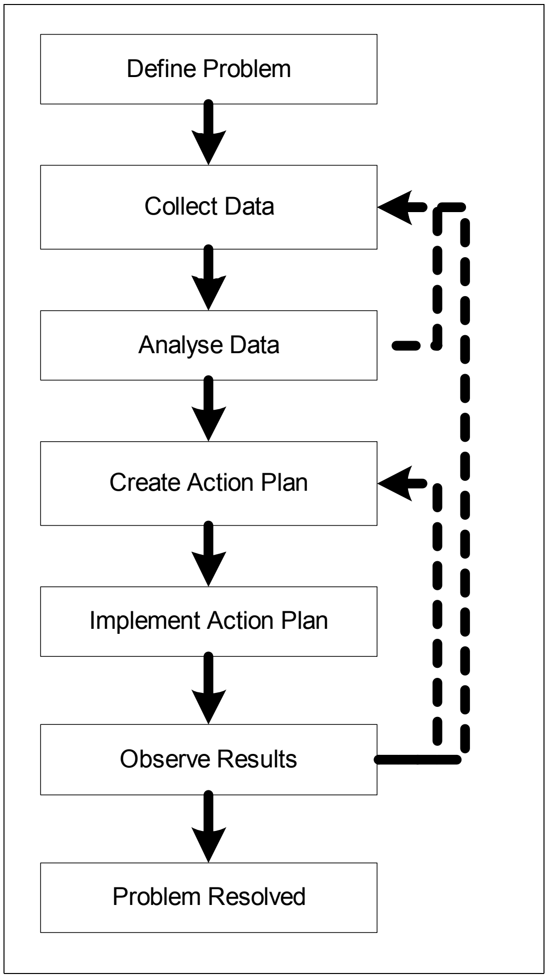


Figure 9-1 Basic flow diagram

9.3.1 Problem definition

In this step, we make an effort to define our problem by asking the following questions:

- Who is affected?
- What are the symptoms of the problem?
- What problems could cause these symptoms?
- When does the problem occur?
- Where in the network are the symptoms seen?
- Is it a hardware or software issue?

9.3.2 Data collection

In this particular step, it might be necessary to collect additional information to help isolate problem cause as noted in the following sections.

Test connectivity

- Use connectivity testing tools such as ping, traceroute, and Telnet
- Note error messages that appear

Details from affected users

Ask more detailed questions:

- Has anything changed recently?
- How often is the problem occurring?
- Is the problem readily reproduceable?
- Is it occurring intermittently or at regular intervals?

Nortel GbESM (Alteon OS) switch dumps

Collect the following dump types:

- TechSupport dump
  - /maint/tsdmp
  - Includes output from the following dumps:
    - Configuration dump (/cfg/dump)
    - Information dump (/info/dump)
    - Statistics dump (/stats/dump)
- Panic dump
  - Manually force a panic if switch hangs
    - /maint/panic
    - <CTRL-SHIFT-"-">
  - /maint/uudmp
  - Includes a core dump of memory registers

  > **Note:** If file size != 65 KB, then the panic file is no good.

- Trace buffers
  - Display switch processor trace buffer contents

    ```
    /maint/debug/sptb <port#>
    ```

  - Display the management processor trace buffer contents

    ```
    /maint/debug/tbuf <port#>
    ```

  - Display the snap trace buffer contents

    ```
    /maint/debug/snap <port#>
    ```

  - Send traces to GbE Technical Support as requested
- Port diagnostics

  Perform a port burn-in test to check for switch port hardware problems. You may also consider running port diagnostics from maintenance kernel.

Network logs

Review the network logs to help you determine possible errors:

- Console log messages

  Last 10 messages only

- Syslog messages

  Highly recommended to use

- SNMP trap messages

  Highly recommended to use

- Server logs

Packet traces

Collect packet traces as needed:

- Client side
- Server side

9.3.3 Data analysis

You should consider specific causes and narrow down the list to the most likely issue or problem. If necessary, gather more data based on your analysis.

9.3.4 Action plan creation

When you have determined the cause of your problem, you should ask the following:

- What are you going to do to resolve the problem?
- Who will do the work?
- When will the work be done?
- Can configuration changes be made during core business hours or is a maintenance window required?
- Who needs to approve the change?
- What is the back-out plan if the problem is not resolved or something goes wrong with the change?

9.3.5 Action plan implementation

Make the planned changes to the environment in an attempt to resolve the problem.

9.3.6 Observation of results

Review your efforts to resolve the problem:

- Did the changes resolve the problem? If not, then you may need to go back a few steps and gather more information or formulate a new action plan.
- Verify that the problem has been resolved. If it has not been resolved, collect more data, analyze, and revise implementation plan.
- Ensure that your changes have not inadvertently caused other problems. If another problem is created, try to troubleshoot that problem (if you have time and approval). Otherwise, initiate the back-out plan.

9.3.7 Problem resolution

Document the problem and resolution for future reference.

9.4 Troubleshooting tools

This section lists troubleshooting tools for your use.

Ping

This tool, also known as the Packet INternet Groper, allows you to check connectivity between network devices. This tool is included with Windows, all variants of UNIX, and is also available on many network device CLIs.

Traceroute

Traceroute records the route through the Internet between your computer and a specified destination computer. If there is a loss of connectivity between the devices, this tool may help in identifying the location of the break in connectivity. This tool is included with Windows, all variants of UNIX, and is also available on many network device CLIs.

Telnet/SSH

Telnet/SSH allows you to connect to device command line interfaces. It is also useful in checking whether a device is listening on a specific port, for example:

```
> telnet 192.168.1.1 80
```

Telnet is included with Windows, all variants of UNIX, and is also available on many network device CLIs.

- Tools for Windows are:
  - Telnet
  - TeraTerm
  - SecureCRT
- Tools for Linux are:
  - Telnet
  - SSH

> **Note:** A basic Telnet client is included with Windows.
>
> TeraTerm (http://www.tucows.com) is a freeware Telnet client and serial communications application. Thus, you can also use it to configure a network device using the serial port of the client station.
>
> SecureCRT is a commercial Telnet and SSH client.
>
> Linux includes text-based Telnet and SSH clients.

Difference tool

The Difference tool allows you to compare the content of two files. For example, if you have two configuration files for a pair of active-standby switches, it helps identify any configuration discrepancies. There are text-based and visual versions for Windows. A text-based version is included with all variants of UNIX.

- Tools for Windows
  - Examdiff
  - VDiff32

> **Note:** ExamDiff is a freeware difference tool that you can download at:
>
> `http://www.prestosoft.com/ps.asp?page=edp_examdiff`
>
> ExamDiff Pro has some enhancements. However, it is shareware. VDiff32 is included with the MKS Toolkit.

- Tool available for Linux
  - diff

TFTP server

A TFTP server is the preferred method for upgrading software images and uploading system panic dumps. You can also use it to upload and download configuration files to and from a network device.

Tools available for Linux and Windows:

- 3Com 3CDaemon (TFTP/FTP/Syslog)
- TFTPD32

> **Note:** The 3CDaemon - 3Com Software Library is a freeware tool that runs on Windows. It also includes TFTP client, FTP server, and Syslog server functionality. This tool is available at:
>
> `http://support.3com.com/software/utilities_for_windows_32_bit.htm`
>
> TFTP is a freeware TFTP server that runs on Windows. It has a very small footprint at only 56 KB. The tool is available at:
>
> `http://www.myzips.com/software/TFTP.phtml`
>
> Linux TFTP server is included with many Linux distributions.
>
> Note that some TFTP servers do not work properly when they are used to upgrade an GbESM firmware. Those listed here have been tested and verified to work properly.

Syslog server

A syslog messaging server for application, network, and operating-system logging is available for Linux and Windows. That tool is 3Com 3CDaemon (TFTP/FTP/Syslog).

Web load generator

A Web load generator allows you to test and analyze the performance characteristics and bottlenecks of your Web site under various load conditions.

- Tools available for Windows:
  - Microsoft Web Stress Tool
  - WebBench
  - LoadRunner
  - Socrates
  - WAPT

> **Note:** Microsoft Web Stress Tool is a free load generator.
>
> WebBench is a free load generator and performance benchmark tool.
>
> LoadRunner is a commercial load generator and performance testing tool that is available at:
>
> `http://www-heva.mercuryinteractive.com/products/loadrunner/`

- Tools available for Linux:
  - Apache Bench
  - Web polygraph

> **Note:** Apache Bench is a freeware load generator that is included with Apache Web server software for Linux.

Protocol analyzer

The Protocol analyzer allows you to take packet captures and read traces that are sent to you by other users.

- Tools available for Windows:
  - Ethereal
  - Etherpeak
  - Sniffer Pro/Net X-ray
  - Shomiti
- Tools available for Linux (UNIX):
  - Ethereal
  - TCPDump
  - Snoop

> **Note:** Ethereal is a freeware tool that runs on Windows and most UNIX variants. It has the ability to read many capture formats including Sniffer, Sniffer Pro, Snoop, TCPdump. The tool is available at:
>
> `http://www.ethereal.com/`
>
> Etherpeek, Sniffer, and Shomiti are commercial protocol analyzers.
>
> Snoop is included with Solaris.

Network monitors or management systems

Use network monitor or management systems to create traffic profiles and to form baseline statistics. Such a system provides notification of network device errors and failures. Brand examples of such systems include Nortel Networks Optivity NMS (network management server), HP OpenView, IBM Tivoli, Ciscoworks, and Concord Network Health.

GbE Switch documents

Here are some GbE documents available for use when troubleshooting:

- GbE Switch Command Reference
- GbE Switch Application Guide
- GbE Switch Module Training Course Labs and Notes

GbE Switch software images

It is good practice to run the most current version of the switch software unless there is a compelling reason not to. The release loaded on a new switch when you receive it might not be the most current version. There are two kinds of software releases:

- GbE Switch feature release, a feature release that adds new functionality and sometimes new commands
- GbE Switch patch release, a patch release that is available as bug fixes are created

10

Service and support

This chapter explains the support methods that you can use if you have issues with the Nortel Networks L2/3 GbESM.

10.1 Placing a call to IBM

For U.S., AP, CAN, and EMEA, use one of the following numbers when calling IBM for technical support:

- Within the United States, call the IBM Support Center at 1-800-IBM-SERV (426-7378).
- Within Canada:
  - For support, call HelpPC at 800-426-7378.
  - For more information or to place an order, call 800-465-7999.
- Outside the United States and Canada, contact your IBM HelpWare® number, your place of purchase, or your local IBM office.

For LA, for technical support, call the IBM HelpCenter®, contact your IBM HelpWare number, your place of purchase, or your local IBM office.

10.2 Online services

For online services for U.S., AP, CAN, and EMEA, visit the following Web site:

`http://www.ibm.com/support/us/`

For online services for LA, visit the following Web site:

`http://www.ibm.com/pc/la`

For online directory services, access the Directory of World Wide Contacts at the following Web site and select your country:

`http://www.ibm.com/planetwide/`

On this Web site, look for the appropriate telephone number under *technical support* and call IBM for assistance.

10.3 Ordering information

The Nortel Networks Layer 2/3 GbE Switch Modules ordering part number is 26K6530 (copper) and 26K6531 (fiber).

- Within the U.S.:

  For information about ordering through PartnerLink®, call 800-426-7272, Option 8. For further details, contact the IBM Remarketer Fulfillment Center at 800-426-9735 or your local marketing support representative.

- Within the EMEA:

  Orders can be entered into the Fulfillment system now. Orders will be addressed for scheduling sequentially. Orders involving multiple units may be subject to an extended delivery schedule. No delivery commitments may be made until schedule is committed. Fulfillment of this product for Personal Computing Division Business Partners is through the SAP/Direct Ship order entry systems and processes.

- Online:

  This product is available online through the BladeCenter Switch Modules Web site:

  `http://www.ibm.com/servers/eserver/bladecenter/switch/more_info.html`

10.4 Other support sites

Listed here are other helpful Web sites:

- Nortel Networks Technical Support

  `http://www130.nortelnetworks.com/cgi-bin/eserv/cs/main.jsp`

- Nortel Networks Products

  `http://products.nortel.com/go/product_index.jsp`

Abbreviations and acronyms

AIX	IBM open operating system
AMD	Manufacturer of personal computing microprocessors
ANSI	American National Standards Institute
ARP	Address Resolution Protocol
ATI	World's largest supplier of 3D graphics and multimedia technology
BACS	Broadcom Advanced Control Suite
BASP	Broadcom Advanced Services Program
BBI	Browser Based Interface
BGPv4	Border Gateway Protocol version 4
BPDU	Bridge Protocol Data Unit
BSMP	blade system management processor
CD-ROM	Compact Disc, read-only-memory
CIOB-X2	Champion I/O Bridge
CLI	Command-line interface
CMIC	Champion Memory and I/O Controller
CSB5	Champion South Bridge
DDR	double data rate
DDR-SDRAM	double data rate-synchronous dynamic RAM
DHCP	Dynamic Host Configuration Protocol
DTP	dynamic trunk protocol
EEPROM	electrically erasable programmable read-only memory
EIGRP	Enhanced Interior Gateway Routing Protocol
ERP	Enterprise Resource Planning
FDD	Floppy Disk Drives
FTP	File Transfer Protocol
GBIC	Gigabit Interface Converter
GEA	Gigabit Ethernet Aggregation
HA	High Availability
HSRP	Hot Standby Router Protocol
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol over Secure Socket Layer
I2C	Inter-IC
IBM	International Business Machines Corporation
ICMP	Internet Control Message Protocol
IDE	Integrated Drive Electronics
IEEE	Institute of Electrical and Electronic Engineering
IEEE 801.1w	Rapid Spanning Tree
IEEE 802.1D	Spanning Tree Protocol (STP)
IEEE 802.1p	Differentiated Service Code Point
IEEE 802.1Q	Virtual Local Area Network (VLAN)
IEEE 802.1s	Multiple Spanning Tree
IEEE 802.3	Link Aggregation Control
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IMB	Inter Module Buses
IMB2	Inter Module Buses2
IOS	Internetwork Operating Software
ISMP	integrated system management processor
ITSO	International Technical Support Organization
LACP	Link Aggregation Control Protocol
LC	
MAC	Media Access Control Address
MIB	Management Information Base
MISTP	Multiple Instance Spanning Tree Protocol
MMF	multimode fiber
MRST	
MST	Multiple Spanning Tree
MSTP	Multiple Spanning Tree Protocol
MTU	Maximum Transmission Unit
NIC	network interface card
NMS	Nortel network management server
OEM	original equipment manufacturer
OSPFv2	Open Shortest Path First version 2
PHY	physical layer transceiver
POST	Power On Self Test

PVID	Port VLAN ID
PVST	Per VLAN Spanning Tree
PXE	Preboot Execution Environment
QOS	Quality of Service
RADIUS	Remote Authentication and Remote Authentication Dial-in User Service Protocol
RIPv1	Routing Information Protocol version 1
RIPv2	Routing Information Protocol version 2
SDRAM	synchronous dynamic RAM
SERDES	serializer/deserializer is an integrated circuit
SLB	Smart Load Balancing
SMP	symmetric multiprocessing
SNMPv3	Simple Network Management Protocol version 3
SoL	Serial over LAN
SSHv2	Secure Shell version 2
STG	Spanning Tree Groups
STP	Spanning Tree Protocol
TACACS	Terminal Access Controller Access Control System
TCP/IP	Transmission Control Protocol/Internet Protocol
TFTP	Trivial File Transfer Protocol
UTP	Unshielded Twisted Pair
VPD	vital product data
VRRP	Virtual Router Redundancy Protocol
WAPT	load and stress testing tool

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this Redpaper.

IBM Redbooks

For information about ordering these publications, see “How to get IBM Redbooks” on page 173. Note that some of the documents referenced here may be available in softcopy only.

- *IBM @server BladeCenter Layer 2-7 Network Switching*, REDP-3755-00)
- *IBM @server BladeCenter Networking Options,* REDP-3660-00

Other publications

These publications are also relevant as further information sources:

- Alteon OS 21.x Application Guide
- Alteon OS 21.x Browser-Based Interface Quick Guide
- Alteon OS 21.x Command Reference
- IBM @server BladeCenter Management Module User’s Guide
- IBM @server BladeCenter Management Module Installation Guide
- IBM @server BladeCenter Planning and Installation Guide

Online resources

These Web sites and URLs are also relevant as further information sources:

- IBM @server BladeCenter network switching

  http://www.bladeserverswitching.org/
- IBM products

  http://www.ibm.com/products/us/
- BladeCenter hardware

  http://www.ibm.com/servers/eserver/bladecenter/index.html
- BladeCenter storage

  http://www.pc.ibm.com/us/eserver/xseries/storage.html
- Advanced server management

  http://www-1.ibm.com/servers/eserver/xseries/systems_management/xseries_sm.html
- BladeCenter deployment

  http://www.ibm.com/servers/eserver/xseries/systems_management/xseries_sm.html

- BladeCenter Management Module firmware

  http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-54939
- *IBM* @server BladeCenter firmware

  http://www-1.ibm.com/servers/eserver/support/bladecenter/index.html
- IBM Support

  http://www.ibm.com/support/
- IBM UpdateExpress

  http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-53046
- Support for IBM xSeries products

  http://www.ibm.com/servers/eserver/support/xseries/index.html
- Broadcom NetXtreme Gigabit Ethernet drivers

  http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-43815
- Cisco Systems Intelligent Gigabit Ethernet Switch Module for IBM @server BladeCenter

  http://www.redbooks.ibm.com/redpapers/pdfs/redp3869.pdf
- Broadcom Advanced Control Suite (BACS)

  http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-43815
- IBM Driver matrixes

  http://www-307.ibm.com/pc/support/site.wss/DRVR-MATRIX.html
- IEEE Standards Organization

  http://standards.ieee.org/getieee802
- Internet Engineering Task Force (IETF)

  http://www.ietf.org
- BCM570X Broadcom NetXtreme Gigabit Ethernet Teaming white paper

  http://www.broadcom.com/collateral/wp/570X-WP100-R.pdf
- Serial over LAN Setup Guide - IBM @server BladeCenter

  http://www-307.ibm.com/pc/support/site.wss/document.do?Indocid=MIGR-54666
- Tucows

  http://www.tucows.com
- Examdiff

  http://www.prestosoft.com/ps.asp?page=edp_examdiff
- 3Com Software Library

  http://support.3com.com/software/utilities_for_windows_32_bit.htm
- TFTP Download

  http://www.myzips.com/software/TFTP.phtml
- Mercury Loadrunner

  http://www-heva.mercuryinteractive.com/products/loadrunner/
- Network protocol analyzer

  http://www.ethereal.com/
- BladeCenter switch modules information

  http://www.ibm.com/servers/eserver/bladecenter/switch/more_info.html

- IBM Directory of worldwide contacts

  http://www.ibm.com/planetwide/

- IBM Latin America

  http://www.ibm.com/pc/la

- Nortel Networks Technical Support

  http://www130.nortelnetworks.com/cgi-bin/eserv/cs/main.jsp

- Nortel Networks Products

  http://products.nortel.com/go/product_index.jsp

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Nortel Networks L2/3 Ethernet Switch Module for IBM eserver BladeCenter

Redpaper

Full Layer 2 switching and Layer 3 routing

Six external multimode fiber or copper GbE interfaces

Hot pluggable switch modules

This IBM Redpaper positions the Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules for IBM eserver BladeCenter and describes how its integrated switch options enable the consolidation of full Layer 2-3 LAN switching and routing capabilities. The Nortel Networks switch modules also provide an upgrade path to full Layer 4-7 services by including 4-7 switch intelligence.

This Redpaper serves as a Best Practices guide for implementing, configuring, and managing Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules for several network topologies. Our topology examples include Nortel Networks, Cisco Systems, and Extreme Networks network environments.

This Redpaper can help you to understand the Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules architecture. It demonstrates how to use specific tools to manage and administer switch module tasks. It also discusses the differences between Nortel Networks and Cisco Systems terminology.

The audience for this Redpaper is experienced systems and network administrators who want to integrate the Nortel Networks Layer 2/3 Fiber and Copper GbE Switch Modules successfully into new and existing networks.